

**THERAPY FOR WORD
FINDING IN APHASIA**
**Effects on picture naming and
conversation**

Ruth Elizabeth Herbert

University College London

Thesis submitted in fulfilment
of the degree of PhD

UMI Number: U602636

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U602636

Published by ProQuest LLC 2014. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ABSTRACT

Therapy for word finding deficits in aphasia have taken two forms: semantic and phonological, with relatively more examples of the former in the literature. Criticisms levelled against such therapies focus on the fact that in most reported cases treatment effects are limited to treated items, and there is very little evidence of real functional change in terms of improvement in everyday speech for the person with aphasia.

Behaviour in conversation can vary and for this reason it is important to establish reliability and stability of the aspects of conversation under scrutiny. This was carried out in the work reported here in order to identify aspects of conversation which might be used as outcome measures for therapy. The analysis of inter and intra-rater reliability and of test retest stability produced a measure which was used to identify the effects of two forms of therapy.

The two forms of therapy were presented consecutively to three people with aphasia. In the first phase phonological and orthographic cues were used. In the second phase participants were encouraged to use the set of treatment words in speech situations, ranging from naming to definition to use in conversation. The effect of each form of therapy on picture naming and on conversation was measured.

The results showed a positive effect of the phonological and orthographic cues for two of the participants in terms of gains in picture naming. For the third participant this therapy was ineffective. The second phase of therapy was effective for all three in terms of gains in items *only* treated in that phase of therapy. The analysis of the conversation data showed unstable baselines for a number of aspects for all three participants. Nevertheless there were some aspects which were stable for a given individual and some evidence of positive changes after therapy.

ACKNOWLEDGEMENTS

I would like to thank most particularly the three people with aphasia, and their significant others, who took part in the study with enthusiasm and interest. It was a real pleasure to work with them and to discuss aphasia and therapy with those who know most about it.

In completing the thesis the advice, support, and wisdom of my two supervisors, Wendy Best and Shula Chiat, have been invaluable. Both Wendy and Shula read with enthusiasm the miles of script that came their way. The meetings we had throughout this time were always interesting and stimulating. I regret not insisting on more meals out but there is always tomorrow ...

While collecting the data for the study I was employed on a grant from the Tavistock Trust for Aphasia. During the data collection period I worked closely with the ideas of Felicity Osborne, and with the physical reality of Wendy Best and Julie Hickin. Those were very fruitful times which I look back on with fondness. Thanks to all three for those happy times. David Howard provided many of the test materials used here, computer programmes to ensure easy data management, and invaluable statistical advice throughout. Thank you.

I would like to thank also staff at University College London for help and support. Jane Maxim, for enthusiastically embracing anything new and challenging, Chris Donlan for humour and advice on statistics, Katie Kay and Caroline Newton for being sane and friendly and keeping things in perspective. At Sheffield Marcin Szczerbinski offered calm advice that was easy to follow.

To all other friends and colleagues who have offered help when needed: Julie (always on the end of the phone when needed), Caroline, Jo, Sarah, Gill, Stephanie for being understanding about the time it has taken and for ongoing friendship. Finally I wish to thank the whole of my extensive and colourful family for being so encouraging throughout, in particular my mum and dad who are very special.

LIST OF CONTENTS

CHAPTER ONE: SPOKEN WORD PRODUCTION IN APHASIA	21
1.0 Introduction	21
1.0.1 Aims and hypotheses of the research	21
1.0.1.1 Conversation and therapy	21
1.0.1.2 Therapy for word finding deficits	21
1.0.1.3 Relating deficit to therapy outcome	22
1.0.2 Overview of Chapter One	23
1.1 Anomia	23
1.1.1 Word finding in aphasia	23
1.1.2 Aphasia syndromes and anomia	23
1.1.3 Anomia: a cognitive approach to assessment	25
1.2 Methodological issues in aphasia research	26
1.2.1 Assessment methods	26
1.2.2 Issues of validity in testing word finding through picture naming	27
1.2.3 Reliability of methods and analyses in aphasia research	28
1.2.3.1 Age related controls	28
1.2.3.2 Error types and error classification	29
1.2.3.2.1 Whole word errors: semantic relationships	29
1.2.3.2.2 Whole word errors: syntactic relationships	29
1.2.3.2.3 Whole word errors: perseverations	30
1.2.3.2.4 Morphological errors	31
1.2.3.2.5 Formal or phonological errors	31
1.2.3.2.6 Mixed errors	32
1.2.3.2.7 Connected speech: circumlocutions	33
1.2.3.2.8 Summary	33
1.2.3.3 Reporting data	35
1.3 Anomia subtypes	37
1.3.1 Semantic deficits and semantic errors	37
1.3.1.1 Background	37
1.3.1.2 Semantic output errors in aphasia	39
1.3.1.3 Relationship between input and output semantic errors	39
1.3.1.4 Single cases with a lexical semantic deficit	40
1.3.1.5 Semantic errors without a lexical semantic deficit	42
1.3.1.6 Variables and semantic errors	42
1.3.1.7 Comparing targets and semantic errors	43
1.3.1.8 Summary	44
1.3.2 Pure anomia	45
1.3.2.1 Classical anomic aphasia	45
1.3.2.2 Pure anomia: a breakdown between semantics and phonology	45
1.3.2.3 Tip of the tongue states in aphasia	46
1.3.3 Phonological deficits	47
1.3.3.1 Classical accounts of phonological impairment	47
1.3.3.2 Single case studies	47
1.3.4 Summary	49
1.4 Variables influencing naming success	50
1.4.1 Semantic variables	50
1.4.1.1 Imageability/Concreteness	50
1.4.1.3 Operativity	51

1.4.2 Lexical variables	52
1.4.2.1 Frequency	52
1.4.2.2 Age of acquisition	53
1.4.2.3 Familiarity	54
1.4.3 Variables acting upon phonological encoding	54
1.4.3.1 Syllable and phoneme length	54
1.4.3.2 Group studies and single case studies	55
1.4.4 Summary	56
1.5 Summary of the Chapter	56
 CHAPTER TWO: MODELS OF SPOKEN WORD PRODUCTION	 57
2.0 Introduction	57
2.1 Methodology	57
2.1.1 Speech errors	57
2.1.2 Reaction time experiments	58
2.1.3 Cognitive neuropsychology	58
2.1.4 Connectionism	58
2.2 Theories of spoken word production	59
2.2.1 Architecture: Boxes, arrows, units and interconnections	59
2.2.2 Architecture: Semantic and phonological levels	59
2.2.3 Evidence for independent levels: semantics and phonology	60
2.2.3.1 Speech errors in normal speakers	60
2.2.3.2 Tip of the tongue states	62
2.2.3.3 Reaction time experiments	62
2.2.3.4 Evidence from aphasia	63
2.2.4 One-step models	64
2.2.5 Two-step models	64
2.2.5.1 Lemma, or word level	64
2.2.5.2 The independence of syntax (lemma) and phonology: tip of the tongue studies	65
2.2.5.3 The independence of syntax (lemma) and phonology: Evidence from neuropsychological studies	66
2.2.5.4 The organisation of syntactic and phonological knowledge	67
2.2.6 Activation within one and two step models	67
2.2.7 Four models of spoken word production	68
2.2.7.1 WEAVER ++ (Levelt, Roelofs and Meyer, 1999)	68
2.2.7.2 Dell, Schwartz et al (1997): Interactive two-step model	70
2.2.7.3 Rapp and Goldrick's Restricted interactive activation (RIA) model	72
2.2.7.4 Caramazza's (1997) Independent Network model	73
2.3 Levels of representation	75
2.3.1 Semantics	75
2.3.1.1 Two systems or one?	75
2.3.1.2 Semantic representations	76
2.3.2 The lemma or word level	76
2.3.3 Morphological representations	77
2.4 Phonological encoding	77
2.4.1 Slot and filler model	77
2.4.2 Serial ordering in phonological encoding	78
2.4.3 Temporal aspects of phonological encoding	78
2.5 Models of spoken word production and aphasic naming errors	79

2.5.1 Background	79
2.5.2 Non-implemented explanations	79
2.5.2.1 Semantic errors in input and output	79
2.5.2.2 Output semantic errors without input semantic errors	80
2.5.2.3 Evidence for lemma level deficit	81
2.5.2.4 Form related errors	81
2.5.3 Computerised simulations of aphasic naming	82
2.5.3.1 Dell et al (1997)	82
2.5.3.1.1 Problems with this simulation	84
2.5.3.2 Foygel & Dell's (2000) semantic phonological model	86
2.6 Summary	87
 CHAPTER THREE: SPOKEN WORD PRODUCTION IN CONVERSATION	 88
3.0 Introduction	88
3.1 Background	88
3.2 Conversation analysis	89
3.2.1 Sequentiality	90
3.2.2 Turn taking	91
3.2.2.1 Turn construction units and transition relevance points	91
3.2.2.2 Types of turns	93
3.2.3 Repair	94
3.3 Conversation analysis and aphasia	96
3.3.1 Turn taking in aphasia	96
3.3.2 Aphasia and repair	98
3.3.2.1 Causes of breakdown in aphasic conversation	100
3.3.2.2 The nature of repair in aphasic conversation	101
3.3.2.3 Influence of partner	102
3.3.2.4 Types of repair in aphasia	102
3.4 Reliability and the measurement of conversation	103
3.5 Summary	104
 CHAPTER FOUR: RECOVERY AND REHABILITATION	 105
4.0 Introduction	105
4.1 Mechanisms of recovery	105
4.1.1 Cortical plasticity	105
4.1.1.1 Neuronal regeneration and cortical plasticity	105
4.1.1.2 Degree of damage and implications for rehabilitation	106
4.1.2 Connectionist accounts of recovery	107
4.2 Psycholinguistic models and principled therapy	108
4.2.1 Effect of cues on spoken word production	109
4.2.1.1 Evidence from normal participants (priming studies)	109
4.2.1.2 Stimulating word production in aphasia	110
4.2.1.2.1 Phonological cues	110
4.2.1.2.2 Phonological facilitation	111
4.2.1.2.3 Differences between semantic and phonological techniques	112
4.2.1.2.4 Orthographic facilitation	113
4.2.3 Therapy for spoken word production	114
4.2.3.1 Design issues in aphasia therapy studies	114
4.2.3.1.1 Design and analysis	114
4.2.3.1.2 Relating breakdown in naming to therapy method	115

4.2.3.1.3 Generalisation to untreated items	117
4.2.3.1.4 Carryover to other speech situations	118
4.2.3.1.5 Maintenance of therapy effects	119
4.2.3.2 Semantic therapy	120
4.2.3.3 Phonological therapy	121
4.2.3.3.1 Davis and Pring (1991)	121
4.2.3.3.2 Raymer et al (1993)	122
4.2.3.3.3 Hillis and Caramazza (1994)	123
4.2.3.3.4 Greenwald et al (1995)	124
4.2.3.2.5 Miceli, Amitrano, Capasso and Caramazza (1996)	125
4.2.3.2.6 Robson et al (1998)	127
4.2.3.2.7 Fink et al (2002)	128
4.2.3.3.8 Self-cueing phonemic cues	130
4.2.3.4 Orthographic therapy	131
4.2.4 Therapies promoting communicative use of words	132
4.2.5 Conversation-based therapy	133
4.3 Summary and implications	134
 CHAPTER FIVE: DESIGN AND METHODOLOGY	 136
5.0 Introduction	136
5.1 Overview of the study	136
5.1.1 Participants	136
5.1.2 Design	137
5.2 Picture naming assessment and other outcome measures	138
5.2.1 Set details	138
5.2.2 Naming test procedure	139
5.2.3 Response categorisation	139
5.2.4 Written naming	140
5.2.5 Conversation	140
5.2.6 Language control tasks	141
5.3 Language assessments	141
5.3.1 Tests of semantic processing	141
5.3.1.1 Pyramids and Palm Trees	141
5.3.1.2 Word to picture matching	142
5.3.2 Auditory input	142
5.3.3 Auditory short term memory	143
5.3.4 Internal phonology	143
5.3.5 Tests of output phonology	144
5.3.6 Sentence comprehension	144
5.3.7 Cognitive skills	144
5.3.7.1 Non verbal problem solving: Raven's Progressive Matrices	144
5.3.7.2 Visual processing: Object decision	145
5.4 Facilitation of spoken word production	145
5.4.1 Design and procedure	145
5.5 Therapy Phase One	147
5.5.1 Design	147
5.5.1 Therapy Phase One Procedure	148
5.6 Therapy Phase Two	149
5.6.1 Design	149
5.6.2 Phase Two Therapy Procedure	150
5.7 Summary	152

CHAPTER SIX: DEVELOPING A MEASURE OF WORD FINDING IN CONVERSATION	153
6.0 Introduction	153
6.1 Measuring conversation	154
6.2 Overview of the reliability study	155
6.3 Development of the measure	156
6.3.1 Background and sources	156
6.3.2 Type of sample	157
6.3.3 Recording format	157
6.3.4 Length of sample	158
6.4 The conversation measure: draft one	158
6.4.1 Number of topics and topic initiator	159
6.4.2 Number of speech units	159
6.4.3 Number of turns	159
6.4.4 Types of turns	160
6.4.5 Lexical selection	161
6.4.6 Trouble-indicating behaviours: Definitions	162
6.4.7 Repair	163
6.4.7.1 Collaborative repair	163
6.4.7.2 Successful or unsuccessful repair	165
6.4.7.3 Type of repair	165
6.4.8 Summary of draft one	165
6.5 Inter and intra-rater reliability	166
6.5.1 Method	166
6.5.1.1 Participants	166
6.5.1.2 Design	166
6.5.1.3 Procedure	167
6.5.1.4 Transcription	167
6.5.1.5 Analysis	167
6.5.1.6 Item agreement	168
6.5.1.7 Data and analysis	169
6.5.2 Results and discussion of item agreement	170
6.5.3 Variables for draft two	173
6.6 Test-Retest stability	173
6.6.1 Participants	173
6.6.2 Procedure	173
6.6.3 Data analysis	174
6.6.4 Results of the test-retest stability analysis	175
6.6.5 Proportional data	176
6.6.6 Discussion of test-retest stability	177
6.7 Summary	178
 CHAPTER SEVEN: ASSESSMENT RESULTS FOR PH	 179
7.0 Introduction	179
7.0.1 Background details	179
7.1 Picture naming	180
7.1.1 Scores at baseline naming	180
7.1.2 Error analysis	180
7.1.3 Psycholinguistic variables influencing PH's naming	182
7.1.3.1 Matched subsets	183
7.1.3.2 Regression analysis	185
7.1.3.3 Discussion of predictor variables	188

7.1.4 Written picture naming	188
7.1.4.1 Method and Results	188
7.1.4.2 Interpretation of results	188
7.1.5 Phonological and orthographic cueing of picture naming	189
7.1.6 Interpretation of naming data	190
7.2 Semantic processing	190
7.2.1 Pyramids and Palm Trees	190
7.2.2 Word to picture matching	190
7.2.3 Interpretation of results	190
7.3 Auditory discrimination	191
7.3.1 Test performance on minimal pairs	191
7.3.2 Error analysis	191
7.3.3 Interpretation of results	191
7.4 Short term memory	192
7.4.1 Test performance	192
7.4.2 Interpretation of results	192
7.5 Reading aloud and repetition	192
7.5.1 Reading aloud words	192
7.5.2 Reading aloud non-words	193
7.5.3 Repetition of words	193
7.5.4 Repetition of non-words	193
7.5.5 Interpretation of reading and repetition results	194
7.6 Tests of internal phonology	194
7.6.1 Test performance	194
7.6.2 Interpretation of results	195
7.7 Sentence comprehension	195
7.7.1 Test performance	195
7.7.2 Interpretation of results	195
7.8 Non verbal processing	196
7.8.1 Non verbal problem solving	196
7.8.2 Visual processing	196
7.8.3 Interpretation of results	196
7.9 Conversation	196
7.9.1 Turn taking	197
7.9.2 Lexical retrieval	198
7.9.3 Trouble indicating behaviours	198
7.9.4 Repair	198
7.9.5 Conversation measure	199
7.10 Summary of PH's language profile	200
 CHAPTER EIGHT: ASSESSMENT RESULTS FOR SC	 202
8.0 Introduction	202
8.0.1 Background details	202
8.1 Picture naming	203
8.1.1 Scores at baseline naming	203
8.1.2 Error analysis	203
8.1.3 Psycholinguistic variables influencing SC's naming	205
8.1.3.1 Matched subsets	206
8.1.3.2 Regression analysis	207
8.1.3.3 Discussion of predictor variables	209
8.1.4 Written picture naming	209
8.1.4.1 Method and Results	209

8.1.4.2 Interpretation of results	209
8.1.5 Phonological and orthographic cueing of picture naming	210
8.1.6 Interpretation of naming data	210
8.2 Semantic processing	211
8.2.1 Pyramids and Palm Trees	211
8.2.2 Word to picture matching	211
8.2.3 Interpretation of results	212
8.3 Auditory discrimination	212
8.3.1 Test performance on minimal pairs	212
8.3.2 Interpretation of results	213
8.4 Short term memory	213
8.4.1 Test performance	213
8.4.2 Interpretation of results	213
8.5 Reading aloud and repetition	214
8.5.1 Reading aloud words	214
8.5.2 Reading aloud non-words	215
8.5.3 Repetition of words	215
8.5.4 Repetition of non-words	216
8.5.5 Interpretation of reading and repetition results	216
8.6 Tests of internal phonology	217
8.6.1 Test performance	217
8.6.2 Interpretation of results	217
8.7 Sentence comprehension	218
8.7.1 Test performance	218
8.7.2 Interpretation of results	218
8.8 Non verbal processing	219
8.8.1 Non verbal problem solving	219
8.8.2 Visual processing	219
8.8.3 Interpretation of results	219
8.9 Conversation	219
8.9.1 Turn taking	219
8.9.2 Lexical retrieval	221
8.9.3 Trouble indicating behaviours	221
8.9.4 Repair	222
8.9.5 Conversation measure	222
8.9.6 Discussion of SC's conversation data	223
8.10 Summary of SC's language profile	224
CHAPTER NINE: ASSESSMENT RESULTS FOR KR	226
9.0 Introduction	226
9.0.1 Background details	226
9.1 Picture naming	227
9.1.1 Scores at baseline naming	227
9.1.2 Error analysis	227
9.1.3 Psycholinguistic variables influencing KR's naming	229
9.1.3.1 Matched subsets	229
9.1.3.2 Regression analysis	231
9.1.3.3 Discussion of predictor variables	232
9.1.4 Written picture naming	232
9.1.4.1 Method and Results	232
9.1.4.2 Interpretation of results	233
9.1.5 Phonological and orthographic cueing of picture naming	234

9.1.6 Interpretation of naming data	234
9.2 Semantic processing	235
9.2.1 Pyramids and Palm Trees	235
9.2.2 Word to picture matching	235
9.2.3 Interpretation of results	236
9.3 Auditory discrimination	236
9.3.1 Test performance on minimal pairs	236
9.3.2 Error analysis	236
9.3.3 Interpretation of results	237
9.4 Short term memory	237
9.4.1 Test performance	237
9.4.2 Interpretation of results	237
9.5 Reading aloud and repetition	238
9.5.1 Reading aloud words	238
9.5.2 Reading aloud non-words	238
9.5.3 Repetition of words	239
9.5.4 Repetition of non-words	239
9.5.5 Interpretation of reading and repetition results	239
9.6 Tests of internal phonology	240
9.6.1 Test performance	240
9.6.2 Interpretation of results	240
9.7 Sentence comprehension	241
9.7.1 Test performance	241
9.7.2 Interpretation of results	241
9.8 Non verbal processing	241
9.8.1 Non verbal problem solving	241
9.8.2 Visual processing	241
9.8.3 Interpretation of results	242
9.9 Conversation	242
9.9.1 Turn taking	242
9.9.2 Lexical retrieval	243
9.9.3 Trouble indicating behaviours	243
9.9.4 Repair	244
9.9.5 Conversation measure	245
9.10 Summary of KR's language profile	246
 CHAPTER TEN: THERAPY RESULTS	 249
10.1 Design and methods of therapy study	249
10.1.1 Participants	249
10.1.1.1 Participant PH	249
10.1.1.2 Participant SC	250
10.1.1.3 Participant KR	251
10.1.1.4 Summary	252
10.1.2 Method	253
10.1.2.1 Therapy Phase One	253
10.1.2.2 Therapy Phase Two	253
10.1.2.3 Assessments One to Five	253
10.2 Effects of therapy on picture naming	254
10.2.1 Effects of therapy on naming overall	254
10.2.1.1 Participant PH	254
10.2.1.2 Participant SC	255
10.2.1.3 Participant KR	255

10.2.2 Therapy Phase One	256
10.2.2.1 Participant PH	256
10.2.2.1.1 Comparison of treated versus untreated sets	256
10.2.2.1.2 Comparison of phonological and orthographic approaches	257
10.2.2.1.3 Discussion of PH's Phase One results	258
10.2.2.2 Participant SC	258
10.2.2.2.1 Comparison of treated versus untreated sets	258
10.2.2.2.2 Comparison of phonological and orthographic approaches	259
10.2.2.2.3 Discussion of SC's Phase One results	260
10.2.2.3 Participant KR	260
10.2.2.3.1 Comparison of treated versus untreated sets	260
10.2.2.3.2 Comparison of phonological and orthographic approaches	261
10.2.2.3.3 Discussion of KR's Phase One results	261
10.2.3 Therapy Phase Two	262
10.2.3.1 Participant PH	262
10.2.3.1.1 Comparison of treated and untreated sets	262
10.2.3.1.2 Comparison of therapy effect across four subsets	263
10.2.3.1.3 Summary of PH's response to therapy Phase Two	264
10.2.3.2 Participant SC	265
10.2.3.2.1 Comparison of treated and untreated sets	265
10.2.3.2.2 Comparison of therapy effect across four subsets	265
10.2.3.2.3 Analysis of variables across four subsets	266
10.2.3.2.4 Summary of SC's response to therapy Phase Two	267
10.2.3.3 Participant KR	268
10.2.3.3.1 Comparison of treated and untreated sets	268
10.2.3.3.2 Comparison of therapy effect across four subsets	269
10.2.3.3.3 Summary of KR's response to therapy Phase Two	270
10.3 Effects of therapy on Conversation	271
10.3.1 Measuring change after therapy	271
10.3.2 Conversation outcome for PH	278
10.3.2.1 Speech units	278
10.3.2.2 Turn taking	278
10.3.2.3 Lexical retrieval and repair	278
10.3.2.4 Proportional data	278
10.3.2.5 Summary of PH's conversation data	279
10.3.3 Conversation outcome for SC	279
10.3.3.1 Speech units	279
10.3.3.2 Turn taking	279
10.3.3.3 Lexical retrieval and repair	280
10.3.3.4 Proportional data	280
10.3.3.5 Summary of SC's conversation data	280
10.3.4 Conversation outcome for KR	281
10.3.4.1 Speech units	281

10.3.4.2 Turn taking	281
10.3.4.3 Lexical retrieval and repair	282
10.3.4.5 Proportional data	282
10.3.4.6 Summary of KR's conversation data	282
10.3.5 Summary of conversation data	283
10.4 Language control tasks	283
10.4.1 Results	283
10.4.1.1 Results for PH	283
10.4.1.2 Results for SC	285
10.4.1.3 Results for KR	285
10.5 Summary of Chapter Ten	286
 CHAPTER ELEVEN: GENERAL DISCUSSION	 287
11.0 Introduction	287
11.1 The conversation measure	288
11.1.1 Applications of the measure	288
11.1.2 Selection of variables	289
11.1.3 Sampling issues	291
11.1.3.1 Conversation partners	291
11.1.3.2 Sample length	292
11.1.3.3 Topic constraint	292
11.1.3.4 Tape recorded data	293
11.1.4 Inter-rater and intra-rater reliability	294
11.1.4.1 Level of agreement and the Kappa statistic	294
11.1.4.2 Sample length	294
11.1.4.3 Defining terms	295
11.1.4.4 Training	296
11.1.4.5 Summary	296
11.1.5 Test-retest stability	297
11.1.5.1 Data analysis methods	297
11.1.6 Suitability as a clinical outcome measure	298
11.1.7 Future directions	298
11.2 Word finding difficulties in aphasia	299
11.2.1 Classical aphasia syndromes and the data from three people with aphasia	299
11.2.2 Theories of spoken word production and the data from three participants with aphasia	301
11.2.2.1 Levels of processing	302
11.2.2.2 Activation	304
11.2.2.3 Summary	306
11.3 Therapy	308
11.3.1 Phase One Lexical therapy	308
11.3.1.1 Relationship between therapy and level of impairment	308
11.3.1.1.1 PH Phase One	309
11.3.1.1.2 SC Phase One	310
11.3.1.1.3 KR Phase One	311
11.3.1.1.4 Summary	312
11.3.1.2 Components of therapy	312
11.3.1.2.1 Principled stimulation	313
11.3.1.2.2 Amount of stimulation	313
11.3.1.2.3 Awareness of deficits	314
11.3.1.2.4 Arousal levels	314

11.3.2 Phase Two Interactive therapy	315
11.3.2.1 Overall effectiveness of the therapy	315
11.3.2.2 Individual responses	316
11.3.2.3 Response of the sets to therapy	316
11.3.2.4 Distinguishing between Phase One and Phase Two effects	318
11.3.2.5 Maintenance of therapy effects	318
11.3.3 Conversation	319
11.3.3.1 Number of variables and effects of therapy	320
11.3.3.2 Effects of therapy on conversation	320
11.3.3.3 Summary	321
11.4 Main findings of the research	321
11.5 Overall Summary	323
References	326
Appendices	
Appendix 1	340
Appendix 2	341
Appendix 3	342
Appendix 4	344
Appendix 5	345
Appendix 6	346

LIST OF TABLES

CHAPTER ONE

TABLE 1.1: production errors in aphasia according to Goodglass (1993) and Nickels (1997)	34
TABLE 1.2: Original data from Schwartz & Brecher (2000) for 15 patients' formal and semantic errors	35
TABLE 1.3: Reanalysed data from Schwartz & Brecher (2000) for 15 patients' formal and semantic errors	36
TABLE 1.4: KE's performance on the set of semantic tasks	41
TABLE 1.5: Three forms of anomia	49

CHAPTER TWO

TABLE 2.1 Semantic representations in three models of spoken word production	76
TABLE 2.2 Data from patients LB and WR (Dell et al, 1997)	84

CHAPTER FOUR

TABLE 4.1 Results of relevant language tests for RBO and GMA (Miceli et al, 1996)	125
TABLE 4.2 Results of relevant language tests for GF (Robson et al, 1998)	128

CHAPTER FIVE

TABLE 5.1 Design of the study	139
-------------------------------	-----

CHAPTER SIX

TABLE 6.1 Background details of the five participants with aphasia and their conversation partners	166
TABLE 6.2 Values of Kappa recommended by Landis & Koch (1970)	169
TABLE 6.3 Item agreement for inter-rater and intra rater reliability	171
TABLE 6.4 Details of the ten participants with aphasia and their conversation partners for test-retest stability study	174
TABLE 6.5 Values of Pearson's R for all conversation variables at assessments one and two	175
TABLE 6.6 Test-retest stability: proportional data	176

CHAPTER SEVEN

TABLE 7.1 PH's scores in picture naming at assessments one and two	180
TABLE 7.2 PH's Picture naming responses in assessments one and two	181
TABLE 7.3 Examples of PH's naming errors	181
TABLE 7.4 Examples of PH's semantic errors	182
TABLE 7.5: PH distribution of semantic error types in assessments one and two	183
TABLE 7.6 PH analysis of psycholinguistic variables in matched subsets	184
TABLE 7.7 PH raw data numbers correct in matched subsets	184
TABLE 7.8 Correlation matrix for psycholinguistic variables and PH naming data	186
TABLE 7.9 PH's written picture naming: % correct and error types	189
TABLE 7.10 PH Cueing of picture naming	189
TABLE 7.11 Conversation measure: data for PH at assessments one and two	200

CHAPTER EIGHT

TABLE 8.1 SC's scores in picture naming at assessments one and two	203
TABLE 8.2 SC's Picture naming responses in assessments one and two	204
TABLE 8.3: Examples of SC's naming errors	204
TABLE 8.4 Examples of SC's semantic errors	205
TABLE 8.5 distribution of semantic error types in assessments one and two	205
TABLE 8.6 SC analysis of psycholinguistic variables in matched subsets	206
TABLE 8.7 SC numbers correct in matched subsets (raw data)	207
TABLE 8.8 Inter-correlation matrix for the set of 188 items and for SC's naming	208
TABLE 8.9 SC Cueing of picture naming	210
TABLE 8.10 SC's distribution of errors in reading aloud words	214
TABLE 8.11 SC's errors in reading aloud	214
TABLE 8.12 SC repetition of words distribution of error types	215
TABLE 8.13 Conversation measure: data for SC at assessments one and two	223

CHAPTER NINE

TABLE 9.1 KR's scores in picture naming at assessments one and two	227
TABLE 9.2 KR's Picture naming responses in assessments one and two	227
TABLE 9.3 Examples of KR's naming errors	228
TABLE 9.4. Examples of KR's semantic errors	228
TABLE 9.5 Distribution of KR's semantic error types in assessments one and two	228
TABLE 9.6 KR analysis of psycholinguistic variables in matched subsets	229
TABLE 9.7 KR numbers correct in matched subsets	230
TABLE 9.8: Inter-correlation matrix for predictor variables for the naming set for KR	231
TABLE 9.9 KR written picture naming examples of errors	233
TABLE 9.10 KR's written naming error distribution	233
TABLE 9.11 KR Cueing of picture naming	234
TABLE 9.12 Distribution of KR's response types in reading aloud	238
TABLE 9.13 Conversation measure: data for KR at assessments one and two	246

CHAPTER TEN

TABLE 10.1. Assessment results for PH, SC and KR	249
TABLE 10.2 Summary of the three participants	253
TABLE 10.3 PH's scores for phonological and orthographic sets Phase One therapy	257
TABLE 10.4 SC's scores for phonological and orthographic sets Phase One therapy	259
TABLE 10.5 KR's scores for phonological and orthographic sets Phase One therapy	261
TABLE 10.6 PH scores at assessments one to five for the four subsets of items	264
TABLE 10.7: SC scores at assessments one to five for the four subsets of items	266
TABLE 10.8: Values of three variables for the four treatment subsets for SC	267
TABLE 10.9: KR scores at assessments one to five for the four subsets of items	269
TABLE 10.10 Conversation Measure Draft Three	272
TABLE 10.11: Baseline conversation data for group (n= 10) assessments one and two	274

TABLE 10.12 Pre and post therapy conversation data for PH	275
TABLE 10.13: Pre and post therapy conversation data for SC	276
TABLE 10.14 Pre and post therapy conversation data for KR	277
TABLE 10.15 PH language control tasks	284
TABLE 10.16 SC language control tasks	285
TABLE 10.17 KR language control tasks	286

LIST OF FIGURES

CHAPTER ONE

FIGURE 1.1: Route to spoken picture naming based on Patterson & Shewell, 1987) 26

FIGURE 1.2: Model of single word processing, Patterson and Shewell (1987) 38

CHAPTER TWO

FIGURE 2.1 Garrett's (1975, 1980) model of spoken sentence production 61

FIGURE 2.2 WEAVER ++ adapted from Levelt, Roelofs and Meyer (1999) 69

FIGURE 2.3 Interactive activation model of Dell, Schwartz et al (1997) 71

FIGURE 2.4 Caramazza's (1997) Independent Network model 74

CHAPTER FIVE

FIGURE 5.1 Design and procedure for naming facilitation (CV spoken condition) 146

FIGURE 5.2 Phase One Therapy: Allocation of items to therapy and control sets 147

FIGURE 5.3 Procedure for Phase One Therapy Orthographic Condition 149

FIGURE 5.4 Phase Two Therapy - Allocation of items to therapy and control sets 151

CHAPTER TEN

FIGURE 10.1 PH's Naming at Assessments One to Five 254

FIGURE 10.2 SC's Naming at Assessments One to Five 255

FIGURE 10.3 KR's Naming at Assessments One to Five 256

FIGURE 10.4 PH's naming of treated and untreated sets (Phase One) at assessments two and three 257

FIGURE 10.5 SC's naming of treated and untreated sets (Phase One) at assessments two and three 259

FIGURE 10.6 KR's naming of treated and untreated sets (Phase One) at assessments two and three 260

FIGURE 10.7 PH therapy Phase Two treated and untreated sets 263

FIGURE 10.8: SC therapy Phase Two treated and untreated sets 265

FIGURE 10.9 KR therapy Phase Two treated and untreated sets 269

CHAPTER ONE: SPOKEN WORD PRODUCTION IN APHASIA

1.0 INTRODUCTION

1.0.1 Aims and hypotheses of the research

1.0.1.1 Conversation and therapy

Rehabilitative attempts to improve word finding in people with aphasia following stroke have shown a good deal of success when improvement is measured in gains in picture naming. This is encouraging, but the question of whether such gains generalise to everyday speech remains unanswered. In the research reported here an attempt was made to address this issue, by developing a reliable quantitative assessment of word finding in conversation, and by comparing participants' scores prior to and after the administration of therapy. The development of the assessment is described, then data relating to the three participants with aphasia described in this thesis is analysed.

1.0.1.2 Therapy for word finding deficits

There has been growing evidence of the effectiveness of two forms of therapy, semantic and phonological, with reports of the former predominating. There are relatively fewer reports of the latter, yet recent studies suggest that this lexical form of therapy may also be effective. Lexical therapy facilitates production of a target word by the provision of information about the word form, either phonological or orthographic. Previous studies have shown item-specific effects of this form of therapy, with gains shown only in items treated in therapy. In the research reported here further evidence of the effectiveness of this form of therapy is provided. In line with previous studies it was predicted that this therapy would show item-specific effects, and that untreated items would not improve.

Given that it is likely that lexical therapy will have only item-specific effects, thus leading to improved production of only a small number of words, it is unlikely that the effects of such therapy will be seen in everyday conversation. It was therefore predicted that there would be no improvement in word finding in conversation, as measured by the quantitative assessment, after the lexical therapy. A therapy which promotes the use of words in everyday speech may, on the other hand, lead to generalised gains in word finding, affecting both treated and untreated items, and to the improved production of

words in conversation. In this research an interactive form of therapy was trialled, which used targeted sets of words, but which encouraged word retrieval in conversational interactions. It was predicted that this therapy would lead to gains in treated and untreated items as measured by picture naming, and also that word finding in conversation would improve. In the research the effects of the two forms of therapy were compared for three individuals with aphasia, both in terms of picture naming and conversation.

1.0.1.3 Relating deficit to therapy outcome

The third line of enquiry investigated in this research concerns the relationship between a participant's specific psycholinguistic profile, and their response to therapy. This is a crucial area of investigation as such evidence may be used to build a theory of therapy, and ultimately aid in the selection of appropriate therapy for a given individual with aphasia. To this end in-depth assessment data for the three participants with aphasia who took part in this research was gathered, and an attempt was made to relate the pattern of impaired and intact processes to the outcomes of the two therapies. Previous studies have found that lexical therapy is effective for people with good semantic and phonological processing, and a deficit in mapping between the two levels. Thus this form of therapy will be most effective for a person presenting with this profile. It is less easy to predict the probable outcomes of the interactive form of therapy for any given individual. In this therapy semantic, syntactic and phonological information supports word retrieval. It may be that this therapy will be effective for a wider range of functional deficits and therefore for someone for whom a lexical therapy is not effective.

In this thesis the assessment data and the naming performance of three people with aphasia are described. The results of two forms of therapy targeting the deficit in spoken word production are presented. In order to measure the impact of therapy on everyday conversation a quantitative measure of word finding in conversation was devised. Details regarding the reliability and stability of this measure are outlined, and its ability to measure change in conversation is examined.

1.0.2 Overview of Chapter One

In Chapter One research into the nature of the word-finding deficit encountered by people with aphasia is described. The Chapter is organised in four sections. In the first section the history of approaches used in this area is outlined. In the second section methodological issues and problems relating to this research are described. In the third section three forms of anomia are described in terms of a cognitive neuropsychological approach to assessment of the disorder. In the fourth section the influence of psycholinguistic variables on word-finding are described. The aim of this chapter is to identify what is known about anomia, and to highlight key issues which will be taken up in the work carried out for this thesis.

1.1 ANOMIA

1.1.1 Word finding in aphasia

Most people with aphasia have difficulties to some extent across auditory comprehension, verbal expression, reading and writing. One of the significant factors affecting verbal expression is a difficulty in finding words, or anomia. Word finding difficulties are apparent in the everyday conversation of most people with aphasia. Nouns and verbs have received the most attention within research into anomia. This language impairment can have a significant effect on the individual's ability to participate in conversation (Lesser & Milroy, 1995; Perkins Crisp & Walshaw, 1999), to extend existing relationships with family and friends, and to develop meaningful new relationships (see e.g. Parr, Byng & Gilpin, 1997). Everyday activities such as using the telephone, as well as those involving more subtle aspects of language, such as gossiping and giving emotional support, can become extremely difficult for the person with aphasia and lead to their avoidance of communication situations. Assessment and therapy for this form of communication impairment needs therefore to address both the linguistic disturbance and the communicative repercussions.

1.1.2 Aphasia syndromes and anomia

A deficit in word finding is the most ubiquitous symptom of aphasia, occurring in all aphasia syndromes. Benson (1979) identified a number of forms of anomia and related

these to the classical aphasic syndromes. *Word production anomia (motor)* involves effortful non-fluent speech output with phonological distortions, and is associated with Broca's aphasia involving lesions to the left frontal lobe primarily. Goodglass and Wingfield (1997) report that Broca's aphasics rely heavily on nouns in connected speech tasks, often fail to produce the target name and may misname the item or produce a phonemic approximation involving struggle and distortion. *Paraphasic word production anomia* involves fluent output with errors at the phonemic level often leading to the production of unrecognisable words. The person may make repeated attempts at the target (*conduite d'approche*) and this form of anomia is associated with conduction aphasia involving lesions of the peri-sylvian area.

What Benson terms *word selection anomia* involves no other production or comprehension deficits and is the main symptom of the classical syndrome anomic aphasia. Anomia in isolation is comparatively rare. In such cases speech is typically fluent with word finding difficulties, and comprehension is preserved. The person may make semantic errors but rarely makes phonological errors, and is able to give detailed descriptions of items they cannot name (e.g. FR described by Avila, Lambon Ralph et al, 2001; GM and JS described by Lambon Ralph, Sage & Roberts, 2000). They will often readily reject their own semantic errors.

Semantic or nominal anomia involves a difficulty in comprehension of single words presented in both the spoken and the written modalities, and is associated with Wernicke's aphasia with lesions in the temporal lobe. Speech is effortless and fluent but with many word selection errors and paragrammatisms, and auditory comprehension is impaired.

Whilst these efforts to delineate the nature of anomia offer some truth and draw broad distinctions between syndromes of aphasia, with lesions in particular areas tending to cause particular patterns of anomia, within each syndrome significant variations of the disorder exist across individuals. Kohn & Goodglass (1985) investigated aphasic naming errors and found no relationship between error pattern and aphasic syndrome. LeDorze & Nespoulous (1989) analysed naming errors from 20 people with aphasia. Participants were assigned to anomia groups based on their error patterns and this was compared to their aphasia syndrome classification. Again, there was no relationship between error pattern and syndrome.

1.1.3 Anomia: a cognitive approach to assessment

The cognitive approach to aphasia locates language breakdown in models of normal language processing. The route to spoken picture naming from Morton's logogen model is shown in Figure 1.1. In this model, as in most current models of spoken word production, there are discrete representations for semantics and for phonology (see Chapter Two for evidence surrounding this distinction). According to this model naming a picture of an item involves picture recognition, accessing a semantic representation, and accessing a phonological form in the output lexicon. Performance on tests of the word's semantics, and phonological output, can distinguish different patterns of breakdown in the processes involved. In aphasia it is generally assumed that recognition of items is intact and that difficulties naming items arise in semantics or after semantics.

The basic assumption underlying the cognitive approach to aphasia is that of modularity. Marr (1982) used this term to refer to the long-standing finding that specific mental functions (e.g. visual processing) can be impaired while other mental functions (e.g. language processing) remain intact. This is a fairly gross example and since then the quest has been on to find more and more specific dissociations of function. Early findings within language processing centred around reading, for example Marshall & Newcombe, (1973) described JC and ST who could read regular but not irregular words. Within spoken word production dissociations in function have been identified between people who have intact semantic processing (e.g. RGB, Caramazza & Hillis, 1990) and those whose semantics are impaired (e.g. KE, Caramazza & Hillis, 1990); and between functions such as reading and picture naming (e.g. MOS described by Lambon Ralph, Cipolotti & Patterson, 1999, whose performance on reading aloud was 100% compared to 75% on picture naming).

Broadly speaking people with aphasia present with a spoken word production deficit arising in semantics, or after semantics has been successfully accessed. This is in line with Benson's (1979) distinction between 'semantic anomia' and 'word selection anomia'. Similarly Gainotti et al (1986) contrast anomia with and anomia without a lexical comprehension disorder. This broad distinction pertains today, although the complexities of phonological encoding are not addressed by this account. More recent accounts have identified three forms of anomia: Lambon Ralph, Sage & Roberts (2000)

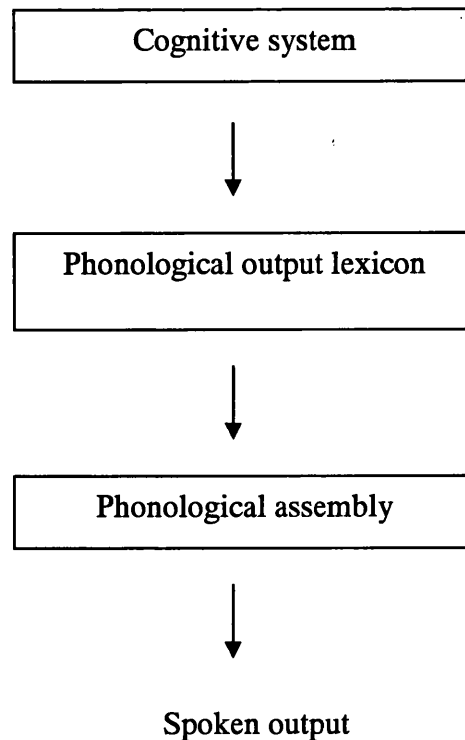


FIGURE 1.1: Route to spoken picture naming based on Patterson & Shewell, 1987)

distinguish between semantic anomia, phonological anomia and classical anomia. The latter group present with word-finding difficulties but without a deficit in semantics or phonology. Within the model shown in Figure 1 the deficit would be located in the mapping from the semantic system to the phonological output lexicon. The next section will highlight the methods used in aphasia research and some of the methodological problems.

1.2 METHODOLOGICAL ISSUES IN APHASIA RESEARCH

1.2.1 Assessment methods

The methods used to attempt to understand word finding deficits have developed over recent years as understanding of the processes involved in spoken word production increases. Picture naming remains the most popular method of measurement. For practical reasons this is warranted: the examiner knows the target, and can compare this with the person's performance; variables pertaining to the target can be manipulated;

reaction times can be measured; scores can be compared across time to investigate recovery or the effects of intervention. Picture naming of nouns and verbs is commonly used in research into aphasia and in clinical investigation of aphasia. The use of pictures necessarily limits the investigation to the study of more imageable items. The method also has more serious limitations. These fall into two categories: issues of validity, and of test reliability. The former will be addressed first.

1.2.2 Issues of validity in testing word finding through picture naming

For a test to be considered valid it must test what it claims to. Critics of the use of picture naming for the assessment of aphasic word finding state that in everyday speech a person is rarely asked to produce a single word. Moreover, unlike in most natural speech situations, picture naming is confrontational in nature. The examinee is aware that there is a desired response. According to some aphasiologists (e.g. Holland, 1994) this method fails to identify the true nature of the problem.

There are however conversational contexts where a single word response is acceptable: the use of ellipsis in responding is a common conversation device (for example in responding to closed questions). Conversation analysis (CA) of normal speakers' everyday talk has dispelled the myth that speech is constructed in grammatically correct sentences and recent work in CA by Helasvuoto (2001) identified interactional roles of noun phrase turns in normal conversation, showing that speakers orient to such turns and co-construct meaning through their use, for example in repair, or in constructing lists.

Evidence from aphasia is equivocal. Crutch and Warrington (2003) report data from FAV who had anomia. Despite a marked difficulty in picture naming FAV's propositional speech was fluent and content-rich. Analysis of connected speech revealed the availability of abstract nouns which, the authors claim, sustained the flow of connected speech. The picture naming task in this instance served to reveal an impairment in noun retrieval that was limited to concrete nouns. Pashek and Tompkins (2002) compared word finding in two tasks: picture naming and a video narration task. They report superior word retrieval in the latter task. However, Berndt, Burton Haendiges and Mitchum (2002) compared speech tasks with and without a picture

present and found no difference across the two types of tasks. These data indicate that both forms of assessment are required.

A further problem concerning the validity of picture naming relates to the fact that bare nouns are being elicited. The phonological form produced in the context of the sentence may differ from that produced in isolation. According to Levelt, Roelofs & Meyer (1999) speech production involves generation of the 'phonological word', where a noun's phonology will vary according to the phrasal or sentential context in which it is produced. The intonational contour may facilitate production of lexical items for people with aphasia (see e.g. melodic intonation therapy described by Sparks, Helm & Albert, 1974). In asking people with aphasia to produce bare nouns in a picture naming test, and therefore depriving them of the support offered by the contextual syntax and intonation, it is possible that performance is artificially lowered. As most studies have used picture naming to measure word finding this problem pertains to most research in the area. It is therefore timely within aphasia research that efforts are made to look beyond single word production tasks. Of interest here is whether picture naming performance matches noun retrieval in everyday speech. Anecdotal accounts of this relationship abound (see, e.g. Goodglass and Wingfield, 1997; Helm-Estabrooks, 1997) and there is a growing interest in research into the relationship.

1.2.3 Reliability of methods and analyses in aphasia research

1.2.3.1 Age related controls

Reliability of an assessment concerns its ability to measure accurately a particular behaviour. On repeat administrations of the assessment one would expect to see similar results for a given subject. The first problem here concerns age-appropriate norms. Many people with aphasia are of older age groups. There is evidence that word finding deteriorates with age in terms of numbers correct (Nicholas, Obler, Albert & Goodglass, 1985) with many researchers finding an increase in tip-of-the-tongue states in older adults (e.g. Burke et al, 1991; Cohen and Faulkner, 1986). In particular the deterioration in performance is found in adults over 70 years of age (Albert, Heller & Milberg, 1988). In a recent study Tsang and Lee (2003) report decreased accuracy and latency in picture naming in older adults (over 70) by comparison with younger adults.

Other researchers have not found a significant difference between older and younger adults: Goulet Ska & Kahn (1994), Van Gorp, Satz, et al (1986) and Nicholas, Brookshire et al (1989) failed to find a significant relationship between age and score on the Boston Naming Test. These conflicting findings indicate that caution is needed in interpreting test results from people with aphasia unless age-matched control data are available.

1.2.3.2 Error types and error classification

People with aphasia make a range of well-documented errors in speech production. Although the classifications used by researchers appear fairly straightforward there are inconsistencies across research teams. This is important as the interpretation of the data is typically used to argue for a particular theoretical position.

1.2.3.2.1 Whole word errors: semantic relationships

For whole word errors the relationship of the error to the target determines the label accorded to the error: calling a monkey an “ape” will usually be labelled a semantic error; saying “Africa” in response to a picture of a monkey may be labelled a semantic error, or a circumlocutory error. Within semantically related errors distinctions between types of errors are made: Coltheart (1980) distinguished between associative errors where the production shares a semantic link with the target (e.g. Africa for target monkey) and co-ordinate errors where the error is a member of the same class of items as the target (monkey and ape). In comparison to conversation where the listener is often ignorant of the target, picture naming allows a more confident appraisal of the relationship between target and production. Even within the constraint of picture naming it is not always clear however what the person with aphasia intended by their response. ‘Africa’ may have been an attempt at a circumlocution (such as ‘lives in Africa’) but sentence processing or verb retrieval deficits prevented this. In such cases subjective interpretation by the examiner is common.

1.2.3.2.2 Whole word errors: syntactic relationships

The error may be syntactically related to the target. Further difficulties arise here. One common but often disregarded problem is that many phonological forms in English correspond to both nouns and verbs, including some very high frequency verbs (e.g. go,

try, take). If the word is produced in isolation it is impossible to determine which syntactic category was intended. The second difficulty was outlined above: if the response to a picture of a monkey were “swing” the examiner, unsure whether a noun or a verb were intended, is unable to determine whether this is an associative error or an attempt at a circumlocution. Again, subjective judgements are common.

Goodglass (1993) categorises errors of syntactic class (e.g. “smoking” for target cigarette) as one-word circumlocutions, assuming that, unable to find the correct target word, the person is attempting to describe something of the concept and manages only one word. This assumption disregards the important issue of shared grammatical class in target and error which is important in theory construction (see Chapter Two).

1.2.3.2.3 Whole word errors: perseverations

Perseveration of a previously uttered sound, word fragment, word or phrase in response to a new stimulus is common in all aphasic syndromes. In severe cases of aphasia with little spontaneous speech recurring stereotypical utterances may predominate, which have been interpreted as resulting from a breakdown in inhibitory mechanisms (Blanken, 1991).

Albert & Sandson (1986) distinguish between repetition of an immediately preceding utterance (which they term ‘stuck in set’ perseveration) and repetition of an earlier not immediately preceding utterance (‘recurrent’ perseveration). This issue of the distance between initial production and recurrence of the item has received some attention.

Martin, Loach, Brecher & Lowery (1998) found that the distance between initial and perseverated response could be lengthy but that perseverations with a semantic relationship to the target could occur over longer time lapses than those with no relationship to the target.

Martin, Gagnon, Schwartz, et al (1996) looked at the semantic errors of a group of 19 people with aphasia. In the initial analysis they compared the occurrence of mixed errors (semantic errors with phonological similarity to the target) with estimates of chance. They then divided the semantic errors into two sets: perseverative semantic errors and non-perseverative semantic errors. Only the non-perseverative semantic errors showed a phonological similarity to the target suggesting a different source for perseverative errors from non-perseverative errors. The authors do not hypothesise the

likely source of the perseverations, beyond referring to working memory. Further research into this is required.

1.2.3.2.4 Morphological errors

Morphological errors remain under debate, as they often involve a target and a response from different syntactic categories and therefore qualify as syntactic errors. For this reason morphological errors pose problems for researchers. Production of the plural noun for the singular in picture naming is a fairly common occurrence in aphasia, but is not considered significant by most researchers or by clinicians. It may be the case that these errors are more significant than has hitherto been thought. This issue remains to be investigated.

1.2.3.2.5 Formal or phonological errors

Formal or phonological errors are errors which are related to the target in terms of their phonology. According to Goodglass and Wingfield (1997) anomic aphasics rarely make phonological errors, which suggests that their difficulty lies in lexical retrieval and not in phonologically encoding a retrieved lexical form. Phonological errors are particularly common in conduction aphasics, who typically make many repeated approximations of the target (see e.g. MB described by Franklin, Buerk and Howard, 2002)

As Nickels (1997) points out, the question of whether formal errors are word substitutions (formal paraphasias), or sound substitutions which happen by chance to be real words, is a significant problem methodologically. One commonly adopted approach to this is to take a word and generate all the possible formal errors by substituting one phoneme (e.g. cat would generate bat, dat, fat etc.). By comparing the incidence of words to non-words in the resulting set one can deduce the likelihood that a phonological error will result in a real word (see Best, 1996; Blanken, 1998 for discussion of these issues).

The degree of phonological similarity required for an error related in sound to qualify as a phonological paraphasia is open to question with different researchers classifying errors along different parameters. For example Dell et al (1997) classify a phonological error as follows: “target and error started or ended with the same phoneme; had a phoneme in common at another corresponding syllable or word position, aligning words

left to right; or had more than one phoneme in common in any position (excluding unstressed vowels)” (Dell et al, 1997: 809). Blanken (1990) however stipulated that errors and targets share a minimum of one sound.

Sound selection errors can also result in non-words. Again, classification is problematic. Goodglass labels these errors ‘phonemic’ or ‘literal’ paraphasias stating that they involve “the production of unintended sounds or syllables in the utterance of a partially recognisable word” (Goodglass, 1993: 79). This definition leaves a lot of room for interpretation, not least the question of how much of the intended target must be present for the error to qualify as a phonemic paraphasia. By exclusion from formal paraphasias, which are sound-related real words, these productions are by definition non-words.

Goodglass’ final category is neologistic paraphasias, which are non-word errors. The example he provides (“tilto” for target table, Goodglass, 1993: 80) raises again the issue of subjective classification of errors, as there is phonological overlap between target and error, and this could be classified as a phonemic paraphasia. He states that ‘the patient’s response ... may be so contaminated with extraneous sounds that the result must be labelled *neologistic*’ (sic) (Goodglass and Wingfield, 1997: 13), which begs the question of how contaminated the response must be.

Nickels (1997) highlights these problems in classification. She labels formal paraphasias “real word responses that are phonologically related to the target... as distinct from the majority of phonologically related responses ... which are non-words” (Nickels, 1997: 140). The issue of word and non-word production is important in theory construction and therefore it is important to distinguish between the two.

1.2.3.2.6 Mixed errors

Whole word errors may arise which are also phonologically related to the target. Goodglass (1993) labels these *phono-semantic blends* and includes those unrelated in meaning to the target (e.g. “cable” for target table), as well as those related in meaning (e.g. “brush” for target broom). The distinction between these two types of errors is an extremely important one that has implications for theory construction. With reference to Dell et al’s (1997) theory of spoken word production the presence of mixed errors at a rate above chance in the corpus of normal speech errors (Dell & Reich, 1981; Martin

Weisberg & Saffran, 1989) and in aphasia (e.g. Martin, Gagnon, Schwartz Dell & Saffran, 1996) provides the rationale for implementing interactive activation within the model. A word that is related in meaning and in sound to the target is more likely to be selected, as it receives activation from both semantic and phonological levels, than one that is related in meaning or in sound alone (see Chapter Two for a full discussion of these issues). Goodglass (1993) terms these errors phono-semantic but only includes errors sharing the same syntactic class as the target. He classified errors which do not share syntactic class with the target as one word circumlocutions.

Within aphasia a number of studies have investigated the occurrence of mixed errors. Martin et al (1996) investigated the phonological relationship between semantic errors and targets in normal speakers and 19 aphasic speakers. They found phonological similarities at levels greater than chance for both groups for first and second phoneme and first stressed syllable.

1.2.3.2.7 Connected speech: circumlocutions

In place of a word a person may produce a description of the item. Fluent anomic speakers typically produce such descriptions when they are unable to access a target word (e.g. GM and JS described by Lambon Ralph, Sage & Roberts, 2000). The accuracy and specificity of the descriptions vary. The description may be accurate and uniquely identifying, for example GM described an acorn as “from a tree, an oak, like a seed” (Lambon Ralph et al, 2000: 188) or it may be extremely vague. Error classification needs to account for this variation. Anomic speakers with grammatically correct fluent output are able to produce phrases and sentences which in some cases describe the target accurately. For agrammatic speakers this is often not possible and their circumlocutions may go unnoticed. Single word utterances may be an attempt at a description. Classification needs to be carefully performed in such cases to capture these attempts.

1.2.3.2.8 Summary

The issue of error classification has been laboured here in order to highlight one of the most pervading difficulties in aphasia research: people with aphasia show great variation, and reports of their aphasia therefore need to be as standardised as possible in order to allow useful comparisons across individuals to be made. Whilst there is more

awareness of these problems there are still difficulties in comparing across individual cases due to examiner bias in analysis. Table 1.1 demonstrates the different approaches taken by Nickels (1997) and Goodglass (1993) to highlight some of the areas of disparity.

As can be seen from Table 1.1 there are errors which, in one or the other classification system, are open to interpretation. Goodglass fails to differentiate between certain important error types, notably in labelling semantic errors from separate syntactic categories circumlocutions, and in failing to distinguish mixed errors which are both semantically and phonologically related to the target, from phonologically related / semantically unrelated real word errors. Nickels on the other hand struggles with the

**TABLE 1.1: production errors in aphasia
according to Goodglass (1993) and Nickels (1997)**

Error type	Target “lobster”	Goodglass (1993)	Nickels (1997)
Meaning related: same syntactic class	crab	verbal paraphasia	semantic shared feature error
Meaning related: different syntactic class	swims	one-word circumlocution	semantic associate error
Sound related real word	obstacle	phono-semantic	phonemic paraphasia <i>or</i> semantic unrelated
Sound and meaning related	oyster	phono-semantic	mixed semantic and phonological
Sound related non- word	lobsil	phonemic paraphasia or neologism	phonemic paraphasia
Phonemically unrelated non-word	nemfag	neologism	neologism

issue of whether sound related real words are word errors or sound errors. In addition, if they are sound errors, they are of the same class of error as sound related non-word errors.

1.2.3.3 Reporting data

A final point in this section concerns the way in which data is reported. A brief example will illustrate this problem. Dell et al (1997) described a computational account of naming and then damaged the model in an attempt to reproduce the overall success rate and error patterns found in data from 21 people with fluent aphasia. The error data were reported as proportions of the set of errors, thus 10 semantic errors out of a total of 100 errors were reported as 10%. Schwartz & Brecher (2000) used predictions from this same model to look at data from a set of 15 people with fluent aphasia. They predicted that certain error types (non-words, formal errors, unrelated errors) would correlate with severity of the naming impairment and that other error types (semantic and mixed errors) would not. In their study the 15 people with aphasia were divided into three groups based on naming success on the Philadelphia Naming Test. Using raw scores the authors then claimed that numbers of non-words, formal errors and unrelated errors

**TABLE 1.2: Original data from Schwartz & Brecher (2000)
for 15 patients' formal and semantic errors**

Patient	% correct (PNT)	Group	No. of formal errors	No. of semantic errors
AF	75%	Mild	5	3
EF	74%	Mild	4	6
GS	70%	Mild	11	4
JB	76%	Mild	2	11
JL	76%	Mild	2	6
MB	71%	Mild	13	11
<i>Mean rank mild</i>			<i>4.1</i>	<i>7.8</i>
AK	42%	Moderate	1	8
CW	59%	Moderate	22	9
DS	51%	Moderate	22	12
<i>Mean rank moderate</i>			<i>8.2</i>	<i>11.0</i>
AH	11%	Severe	23	9
AS	15%	Severe	19	9
EG	3%	Severe	45	5
ES	13%	Severe	25	3
ET	3%	Severe	14	2
WR	8%	Severe	25	10
<i>Mean rank severe</i>			<i>11.8</i>	<i>6.8</i>
Kruskal-Wallis test (DF = 2)			K= 9.0, p= 0.01	K=1.8, p= 0.40

increased as severity increased and that semantic and mixed errors did not co-vary with naming success. In their analysis they reported raw error scores thus neglecting to take

account of the error space available. The formal errors and semantic error data from the 15 patients are shown in Table 1.2.

If however the errors from these subjects are treated as proportional data the analysis reveals an entirely different picture. In Table 1.3 the error data has been reanalysed by computing the number of items correct for each individual, and thus the number of errors. The total of each error type was then computed as a proportion of the error space available. Thus AF scored 75% giving a total correct of 131, and an error space of 44. The formal errors ($n=5$) would be computed as $5/44$ or 11%, the semantic errors as $3/44$ or 7%. The reanalysed data using this method is presented in Table 1.3.

Taking formal errors first, in the original analysis these were deemed to be severity sensitive. In the new analysis the overall result is significant but clearly the percentage of formal errors does not co-vary with severity. The mild and severe groups have equivalent mean ranks. The statistically significant result comes from the moderate group which, due to two members of the group recording a large number of formal errors, has a high mean score of errors and a high mean rank score.

**TABLE 1.3: Reanalysed data from Schwartz & Brecher (2000)
for 15 patients' formal and semantic errors**

Patient	% correct (PNT)	Group	% of formal errors	% of semantic errors
AF	75%	Mild	0.11	0.07
EF	74%	Mild	0.09	0.13
GS	70%	Mild	0.21	0.08
JB	76%	Mild	0.05	0.26
JL	76%	Mild	0.05	0.14
MB	71%	Mild	0.25	0.22
<i>Mean rank mild</i>			6.5	11.25
AK	42%	Moderate	0.04	0.08
CW	59%	Moderate	0.31	0.13
DS	51%	Moderate	0.26	0.14
<i>Mean rank moderate</i>			9.83	10.5
AH	11%	Severe	0.15	0.06
AS	15%	Severe	0.13	0.06
EG	3%	Severe	0.26	0.03
ES	13%	Severe	0.16	0.02
ET	3%	Severe	0.08	0.01
WR	8%	Severe	0.16	0.06
<i>Mean rank severe</i>			6.58	3.5
Kruskal-Wallis test (DF = 2)			K=10.4, p<0.01	K= 21.89, p<0.001

Semantic errors were deemed to be severity insensitive. Again taking the proportion of errors which are semantic the group analysis is significant showing that the proportion of semantic errors decreases as severity of the naming deficit increases. Similar conflicting results are found for the other error types on similar re-analysis of the data. This example serves to illustrate the different interpretations that can be applied to data and the importance of selecting the most appropriate analysis. If two forms of analysis are available both should be examined and inconsistencies between the two highlighted.

1.3 ANOMIA SUBTYPES

In this section the cognitive approach to aphasia is considered and the analysis of anomia in terms of a single word processing model is outlined with reference to recently published single case data.

1.3.1 Semantic deficits and semantic errors

1.3.1.1 Background

As stated above (section 1.1.2) Benson (1979) described semantic or nominal aphasia where the individual presents with a lexical comprehension deficit. Gainotti (1976) made a similar distinction between people with and without a lexical comprehension disorder. In the model shown in Figure 1.1 this would constitute a breakdown in processing within the level of semantics.¹

The model shown in Figure 1.2 includes access routes to semantics from auditory and visual stimuli, and a further route from semantics to written output. Within this theory it is possible to test central semantic representations through input tasks and output tasks. Input tasks of semantics include word to picture matching, picture verification, synonym and antonym judgements, and word association tasks. Items may be presented auditorily or visually. All of these tasks involve the selection of a target item from an array or a choice of two, except for picture verification which involves a yes / no response. In most tasks semantic foils are present which vary in the closeness of their semantic relationship to the target. Thus errors in input tasks can be analysed in terms of

¹ There are a few reported cases of individuals with aphasia who are able to name pictures for which they show no comprehension.

their association to the target and conclusions drawn from the selections made. This is of particular note in cases of progressive loss of semantic representations such as is seen in semantic dementia or in Alzheimer's disease. Here the progression of the disease may be seen in a deterioration in picture naming and in the widening relationship between target and error. Progressive deterioration is not found in neurologically stable people with aphasia, but the relationship between target and error can serve to identify the degree of severity of a semantic impairment with closely related errors being common in less severe cases.

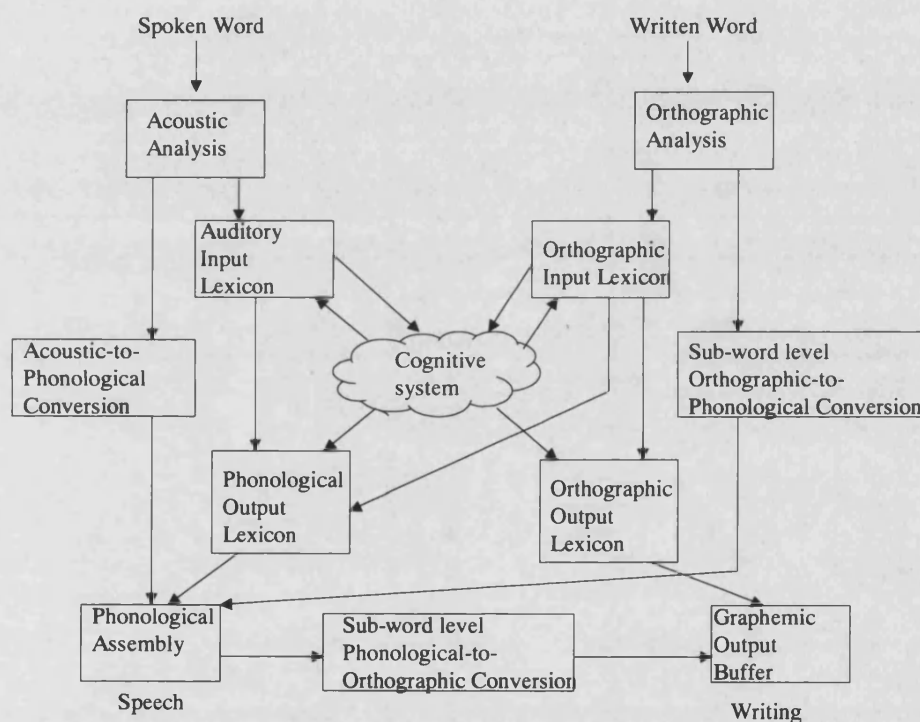


FIGURE 1.2: Model of single word processing, Patterson and Shewell (1987)

All of the tasks described above are off-line tasks. It has been argued by many researchers (e.g. Nickels, 1997; Lambon Ralph et al, 2000) that input tasks are much less demanding than output tasks, as in the latter the correct semantic and lexical representations must be selected from the set of all possible choices. In input tasks activation of semantics may occur via visual input from the set of pictures and via the input phonological form. The summation of these two forms of activation may serve to facilitate the correct selection. Best and Nickels (2000) recommend the use of tasks

where controls are not at ceiling such as reaction time comparisons, to ensure a valid appraisal of aphasic semantic processing.

1.3.1.2 Semantic output errors in aphasia

Semantic output errors are a common occurrence in aphasia. People with aphasia make semantic errors in naming pictures (e.g. KE, described by Caramazza & Hillis, 1990), in reading aloud (e.g. AB and CD, Coltheart, Patterson & Marshall, 1980) and in repetition (e.g. NC described by Martin, Dell, Saffran & Schwartz, 1994). There are reports of people with aphasia who make semantic errors in one modality only, for example in reading but not in speech (e.g. Coltheart, Patterson & Marshall, 1980). All people with aphasia regardless of syndrome make semantic errors in spoken picture naming (Goodglass and Wingfield, 1997). The production of semantic errors has been well researched both in an effort to understand the nature of the deficit, and to attempt rehabilitation of the language impairment.

Early explanations of semantic errors in aphasia (e.g. Rinnert and Whitaker, 1973) proposed that word associations produced by normal speakers derived from the same source as semantic errors in aphasia. Deloche, Hannequin, Dordain et al (1996) have cast doubt upon this contention. They compared naming errors in normal and aphasic speakers and showed different variables affecting normal and aphasic errors and a disparity across the two groups in terms of which items were difficult to name.

Semantically related errors have been differentiated along the lines Coltheart (1980) proposes (see 1.2.3.2.1). Within the subset of shared feature errors further discriminations can be made: errors may be super-ordinate, category co-ordinate, or category subordinate. Examination of semantic errors in aphasia has found evidence of shared feature errors predominating over associative errors, but no evidence of one type of shared feature error being more prevalent than another (see Nickels, 1997).

1.3.1.3 Relationship between input and output semantic errors

According to Gainotti (1976) and Gainotti et al (1981) speakers with aphasia who make more semantic errors in naming also make more semantic errors in comprehension. Gainotti et al (1986) investigated 13 people with dysphasia with naming problems. All

had fluent well articulated grammatically correct speech with few phonemic paraphasias. Lexical comprehension tests (spoken and written word to picture matching with related distractors) were used to detect a deficit. The severity of the anomia was equal across the two groups. Those with a lexical comprehension deficit made more spoken semantic errors than the other group and showed little knowledge of the word form when they were unable to name the picture.

Butterworth, Howard & McLoughlin (1984) found a relationship between the number of semantic errors in comprehension and the number produced in naming for 30 people with aphasia. Similarly Nickels & Howard (1994) found that performance on high imageability single word comprehension tasks predicted the number of semantic errors in production. Butterworth et al (1984) and others (Howard & Orchard-Lisle, 1984; Howard, Patterson et al, 1984) found no item by item consistency for semantic errors. Thus errors are not thought to be caused by the loss of representations for particular items (unlike the deficits seen in people with dementia or semantic dementia), but rather are due to disturbed access routines.

1.3.1.4 Single cases with a lexical semantic deficit

Howard and Orchard-Lisle (1984) described JCU. She was severely aphasic following a stroke. She demonstrated no problems in picture recognition but was severely impaired in naming pictures (3% correct). Phonological cues helped significantly, and when she was provided with a phonological miscue (e.g. target shoe cued with /bə/) she could be miscued into production of a semantic co-ordinate (e.g. boot). When these errors were re-presented to her in a picture verification task (e.g. shown a picture of a tiger and asked "Is it a lion?") she accepted 56% of her own error productions as the correct name for that stimulus picture. The data led Howard and Orchard-Lisle to conclude that JCU was operating with under-specified semantic representations. Her deficit was thought to arise at the semantic level within a one stage model of single word production such as that proposed by Patterson and Shewell (1987) (see Figure 1.2). Here semantics directly address an entry in the phonological output lexicon. Failure at the semantic level will involve difficulties with input tasks and word finding and may be associated with semantic errors in production.

Hillis, Rapp, Romani, & Caramazza (1990) described KE, a 52 year old right handed male who had sustained a thrombo-embolic stroke six months prior to the study. KE's speech output was limited to single nouns, and well-known phrases, and he made semantic errors in speech. Standard tests showed an auditory comprehension deficit. The authors used the same list of items to test all of the following: auditory and visual word to picture matching, oral naming, written naming, oral reading and writing to dictation. The results are shown in Table 1.4.

As KE showed impairment in all six tasks and produced semantic errors in all of the tasks as well, they concluded that KE had a central semantic processing deficit affecting all tasks reliant upon semantic mediation. This case and that of JCU map neatly onto the single word processing model considered so far. The deficit in both cases can be located within semantics.

TABLE 1.4: KE's performance on the set of semantic tasks

Task	Total Correct (n=144)	% correct	Semantic errors/total errors	% errors = semantic
Auditory word to pic matching	83	0.58	58/61	0.95
Visual word to pic matching	91	0.63	39/53	0.74
Oral naming	80	0.56	59/64	0.92
Written naming	77	0.53	50/67	0.75
Oral reading	84	0.58	52/60	0.87
Writing to dictation	84	0.58	40/60	0.67

Alario, Schiller, Domoto et al (2003) investigated the relationship between the status of sub-lexical processing and the production of semantic errors. They report data from two people with aphasia who presented with mild lexical semantic deficits and made semantic errors in picture naming. One of the two had an impairment affecting transcoding tasks such as reading aloud and repetition, and also made semantic errors in these tasks. The other showed intact processing in transcoding tasks, and made no semantic errors in these tasks. The authors claim that the sub-lexical mechanism acts to constrain phonological production. Where sub-lexical routines are operating well, phonological information is available and phonologically unrelated semantic errors will be less likely to be selected than the target word form. Where no or reduced support

from sub-lexical routines exists semantic errors are likely in the presence of a semantic deficit. They term this the 'summation hypothesis'.

1.3.1.5 Semantic errors without a lexical semantic deficit

Caramazza & Hillis (1990) described two people with aphasia who presented with intact semantic processing, as shown on tests of input to semantics such as word-to-picture matching, yet made semantic errors in spoken picture naming. This poses a problem for a model which incorporates only semantics and phonology. It is unclear where the semantic errors in output arise. Loath to add a further level of representation, Caramazza & Hillis suggested that RGB and HW had a deficit in accessing phonological representations: when the correct representation was unavailable they were forced to select the nearest neighbour. This account puts some strain on a theory incorporating one step from meaning to sound, which will be addressed in the discussion of relevant theories in Chapter Two.

1.3.1.6 Variables and semantic errors

If it is the case that specific representations are not lost in aphasia, but that access to semantic representations is disturbed, investigation of the psycholinguistic variables known to affect language processing may shed light on the source of semantic errors.

Nickels & Howard (1994) investigated semantic errors in 15 people with aphasia. They subdivided the group on the basis of their performance on input tasks and on the number of semantic errors they produced. There was a strong relationship between performance on input tasks and the number of semantic errors. They predicted that the semantic group (those who made semantic errors on input and output) would show an effect of imageability on error production, as their errors were thought to arise at the semantic level. As a group the prediction was borne out, with imageability predicting both correct responses, and the occurrence of semantic errors. However, when individual cases were analysed, imageability predicted correct responses and semantic errors for only four of the eight. Imageability did not predict responses for the non-semantic group (those who made no errors for high imageability input tests and made few output semantic errors).

For two of the semantic group there was an interaction between familiarity and imageability. Nickels and Howard explain this as follows: for highly imageable items a correct response is available; for low imageable items a range of items are accessed at the level of phonology, the most high frequency item being selected. Thus high frequency items will be named correctly, low frequency items will result in semantic errors. It is unclear however why only some of the semantic group show an imageability effect if the source of their errors is identical across the group. One possible explanation considered by the authors concerns the nature of the stimuli: being pictureable items these were all highly imageable, thus the range of imageability dealt with in this study is narrow. This research demonstrates the conflicting findings which may emerge when group and individual data for people with aphasia are compared.

Further evidence for the effect of variables on production of errors comes from RGB and HW (Caramazza and Hillis, 1990). These two participants with aphasia were both thought to have a problem in accessing correct phonological representations. Yet for only one of the two (RGB) did frequency predict the occurrence of semantic errors, with mid-frequency items being likely to result in semantic errors. If, as was argued in this paper, the two participants had the same deficit in production both should be affected by the same variables. In unpublished data from a single participant with bilingual aphasia (Herbert, 1998) CL showed a marked effect of familiarity on the production of correct responses and semantic errors, with correct responses being made to high familiarity targets, and semantic errors to mid familiarity targets. Other responses, which were mainly failures to respond, were produced to low familiarity items.

Nickels (1995) looked at the effects of target words' imageability and length on the occurrence of semantic and phonological errors. She found that imageability predicted the production of semantic errors and length the production of phonological errors. Further research is needed to identify the patterns and degree of deficit and the effect of variables on the production of correct responses and semantic errors in order to better inform our understanding of the deficits and to better target rehabilitation efforts.

1.3.1.7 Comparing targets and semantic errors

In the studies referred to above the effect of key variables was investigated in terms of the target item's status. These studies looked at whether variables pertaining to the

target item predict naming success (Caramazza and Hillis, 1990; Nickels and Howard, 1994) and semantic error production (Nickels and Howard, 1994). Few studies have looked at the properties of the *error* words themselves. Evidence from normal speech errors indicates that errors are more likely to be higher in frequency than the intended target, than vice versa (Levelt, 1989). Levelt (1983) and Martin Weisberg and Saffran (1989) failed to replicate this finding however. Hillis and Caramazza (1995) looked at a small corpus of errors from three aphasic participants, each differing in the hypothesised source of their errors. For all three, the mean frequency of the errors was higher than that of the intended targets. No statistical analysis is given and the corpus is extremely small (range 12 – 17 items per person).

Marshall, Pring, Chiat & Robson (2001) report on the reverse frequency effect found in a person with aphasia JP. She showed a preference for low frequency items in all output tasks. Comparison of her semantic errors with their targets revealed that the errors were also much lower in frequency than their corresponding targets. Clearly the predictor variables operating for a given individual need to be considered as potential predictors for error word production.

Thus there is evidence that two distinct patterns of deficit occur in the context of output semantic errors. The degree to which errors in input or output can be predicted by variables remains unclear. Within one set of aphasic speakers presenting with similar symptoms it is probable that marked differences are present and further research is needed to start to identify these differences.

1.3.1.8 Summary

Two sources of semantic errors have been identified in the literature: patients may present with an impairment on tests of input semantics such as word to picture matching and may also make semantic errors in output (e.g. JCU described by Howard & Orchard-Lisle, 1984; KE described by Hillis et al, 1990). Others exist who do not make errors in input tasks, but who still produce semantic errors in speech production (e.g. RGB and HW described by Caramazza and Hillis, 1990). In the former case it has been proposed that the deficit lies in the semantic level of representation. Howard & Orchard-Lisle (1984) argue that JCU is using an incomplete semantic specification in both input and output tasks. In the cases of RGB and HW Caramazza & Hillis (1990) suggest that

the correct lexical semantic representation is accessed, but that the corresponding phonological form is not available and therefore a semantically related neighbour is accessed in its place. It is proposed that the deficit lies in access to or within the phonological output lexicon².

Since RGB and HW were described a number of similar case reports have emerged. These are people with a deficit in accessing phonology from semantics and are labelled nowadays pure anomic. They have intact semantics and intact phonology.

1.3.2 Pure anomia

1.3.2.1 Classical anomic aphasia

According to the classical taxonomy of aphasia, the syndrome anomic aphasia involves no deficit in auditory comprehension, and the person produces fluent well-articulated grammatically correct speech. The main symptom is anomia, which is apparent in spoken output where they will produce errors and may abandon utterances as the words elude them.

1.3.2.2 Pure anomia: a breakdown between semantics and phonology

In pure anomia the person understands the words they are unable to say, as shown by tests of input to semantics. Indeed they are often able to describe the concept in full even though they are unable to produce the word itself. Phonological representations are intact as shown by good performance on tasks such as reading aloud and repetition. The process of translating meaning into sound is therefore thought to be interrupted after semantics and before the phonological representation has been accessed. Errors in spoken word production include semantic errors, which are often rejected, descriptions of the item in question, and no responses. People with pure anomia are often helped to find the word by first sound cues (see e.g. GM and JS described by Lambon Ralph, Sage & Roberts, 2000).

² RGB also had a category specific deficit which Caramazza and Hillis (1990) claim also to arise in the phonological output lexicon, but which later they claim to be due to confounding variables (Caramazza and Hillis, 1995).

This deficit may arise after stroke or other acute aetiology, or may be of a progressive nature. Graham, Patterson & Hodges (1995) describe the progressive aphasia of FM who presented with good lexical comprehension but severe word finding problems. Comprehension remained stable while the anomia progressed considerably over the two year course of the study.

Lambon Ralph et al (2000) describe two cases with pure anomia, GM and JS. Both were within the normal control range for all tests of input to semantics, and both showed good reading aloud and repetition. Both showed impaired picture naming although GM was less impaired than JS. Their responses consisted of semantic errors, all of which were rejected, circumlocutions, and no responses. GM showed some morphological and phonological knowledge of items he was unable to name (e.g. for compound nouns such as seahorse GM was able to identify that this consisted of two other words).

1.3.2.3 Tip of the tongue states in aphasia

Such cases are of interest to theoreticians as they provide the aphasic parallel to the tip-of-the-tongue state seen in normal speakers, where the person knows the word they are trying to say but cannot say it. There have been many studies of this phenomenon in normal speakers looking at how much knowledge of the word is available in the ToT state. The results of such investigations are used to support particular theories in spoken word production (see Chapter Two).

GM and JS (Lambon Ralph et al, 2000) both had difficulties naming pictures. When they were in a ToT state the authors investigated their ability to identify the following: whether a word was a compound noun or not, the number of syllables in the word, and the first letter of the word. Both were able to identify whether the target was a compound noun, GM had access to syllable length information, but neither had access to first letter or sound information.

These cases demonstrate a breakdown in processing between semantics and phonology but suggest that partial phonological information is still available. Both people responded to cues: reading aloud the whole word and hearing the initial sound both facilitated spoken word production. GM was cued by syllabic information and stress

information as well³. Unlike JCU (Howard & Orchard-Lisle, 1984) neither GM nor JS was miscueable.

There are many reports of cases of people with word finding difficulties which investigate the person's knowledge of the word when they are in a ToT state, and a more thorough review of this literature is provided in Chapter Two where it is appraised with reference to theories of spoken word production.

1.3.3 Phonological deficits

1.3.3.1 *Classical accounts of phonological impairment*

Benson's (1979) distinction between *word production anomia* (motor) and *paraphasic word production anomia* is useful here. The former involves effortful non-fluent speech output with phonological distortions, and is associated with Broca's aphasia. The latter involves fluent output with errors at the phonemic level often leading to the production of unrecognisable words. Repeated attempts are made to correct the production (*conduite d'approche*). It is associated with conduction aphasia (see e.g. MB, described by Franklin, Buerk and Howard, 2002). In the case of paraphasic word production anomia it is believed that the person has an accurate semantic representation and has accessed the correct phonology but that there is a breakdown in maintaining the phonological form for the process of phonological encoding. In motor word production anomia the difficulty may lie beyond the phonological stage in translating the phonological representation into a motor or phonetic plan.

1.3.3.2 *Single case studies*

Kay & Ellis (1987) described the much-reported EST. He performed well on tests of input semantics such as synonym judgements, and Pyramids and Palm Trees. Picture naming was impaired (he scored 22% correct on the Boston Naming Test) and was affected by frequency. Phonemic cues did not help except in generating a phonological approximation to the target. Moreover EST did not make semantic errors and could not be miscued to do so. This is quite a rare pattern in aphasia and is unlike the pattern

³ The authors also investigated methods of inhibiting naming: semantically related primes inhibited target production; semantic relatives in blocks inhibited naming as compared to random presentation.

found in normal speech errors where semantic errors are more common than formal errors (e.g. Martin Weisberg & Saffran, 1989). Kay and Ellis proposed that EST's deficit was due to impaired phonological representations. In Benson's terms he presented with paraphasic word production anomia.

NC (Martin & Saffran, 1992; Martin Dell Saffran and Schwartz, 1994) presented with predominantly formal errors and relatively few semantic errors in spoken word production. Blanken (1990) describes a similar pattern. Interestingly NC made semantic and other errors in repetition and could not repeat non-words, a pattern termed deep dysphasia (the parallel to deep dyslexia but involving spoken rather than visual input)⁴. Analysis of NC's formal paraphasias showed a tendency for errors to maintain the initial phoneme of the target (but not the first stressed vowel or final consonant); to preserve the grammatical class of the target; and to be higher in frequency than the target. These data support the contention that the formal paraphasias are 'lexically generated.' The authors go on to propose a pathologically increased decay rate within a model of interactive activation (Dell, 1986; 1988) as the basis of these errors: as the target's activation level subsides rapidly, competitor nodes activated by feedback from the phoneme level to the lexical level are more likely to be selected. These will be related phonologically to the target.

Best (1996) describes the aphasic speech of MF. As she reports formal paraphasias are real words which are related phonologically to the target. Of interest is whether such productions are real words by chance. By comparing MF's errors to a corpus of pseudo-errors generated from MF's own errors, Best (1996) found that the former were less likely to be real words than the latter, suggesting that MF's own errors were formal paraphasias and not real words by chance.

Hillis, Boatman, Hart & Gordon (1999) describe a person with aphasia who made mainly formal errors and no semantic errors in spoken output. JBN was aphasic after a sub-arachnoid haemorrhage. She presented with a fluent aphasia, affecting mainly spoken input and output. Written lexical decision, written picture verification and written naming were intact. In contrast auditory lexical decision and auditory picture

⁴ NC has impaired sublexical mechanism and thus cannot repeat non-words. The presence of semantic errors in his spoken output supports Alario et al's (2003) claim for the summation hypothesis: the lack of phonological information from the sub-lexical route leaves the semantic route unconstrained and thus semantic errors arise.

verification were impaired. Errors in the latter were phonologically related to the target. JBN made formal errors and produced neologistic output in picture naming and oral reading but never made semantic errors. The authors interpret the pattern of results in terms of weakened connections between lexical representations and sub-word phonological units. Weakened connections mean that the correct phonological nodes are not accessed and formal errors occur due to competition from other activated lexical nodes.

Franklin, Buerk and Howard (2002) report data from MB who presented with reproduction conduction aphasia. They describe this as a “phonological impairment of production at the single word level” (Franklin et al, 2002: 1088). The impairment is evident in all speech production tasks and word length is a significant factor. MB produced mainly phonological errors in spoken output, either omitting or substituting phonemes, and had a marked length effect, with longer words being more susceptible.

1.3.4 Summary

TABLE 1.5: Three forms of anomia

Task	Semantic anomia	Pure anomia	Phonological level deficits
Word to picture matching	Impaired	Intact	Intact
Spoken naming	Impaired	Impaired	Impaired
Written naming	Impaired	?	?
Reading aloud	Impaired	Intact	Impaired
Repetition	Impaired	Intact	Impaired
Effect of phonemic cues	Cueable	Cueable	May help
Effect of miscues	Miscueable	No effect	No effect
Errors in spoken naming	Semantic	Semantic – will reject these Circumlocutions Fail to respond	Phonological
Variables affecting naming	Imageability?	?	Frequency? Length

In this section three forms of spoken word production impairment have been described: those involving semantic deficits, pure anomia, and phonological output deficits. Table

1.5 outlines the performance of the three types of anomia on a set of commonly used tests.

1.4 VARIABLES INFLUENCING NAMING SUCCESS

Investigation of the influence of certain variables on aphasic word finding has led to a better understanding of the deficits and their source. These variables split into three categories: those affecting semantic processing, those affecting lexical retrieval and those affecting phonological encoding. It is worth noting at this stage that some variables have confounding effects, for example word frequency and familiarity are known to co-vary thus the independent effects are hard to disentangle.

1.4.1 Semantic variables

1.4.1.1 Imageability/Concreteness

Both the imageability and the concreteness of the referent are commonly used semantic measures. Whilst the terms are often used interchangeably they refer to different properties. Coltheart (1981) used normal participants' ratings of how easy they found it to create an auditory or visual image of the referent to derive imageability ratings. Normal participants' ratings of how accessible to sensory experiences the referent is defines concreteness.

Evidence of an imageability effect is thought to implicate the semantic system. Imageability effects have been found in word finding (e.g. Nickels & Howard, 1994), reading aloud (deep dyslexia, described by e.g. Coltheart, Patterson & Marshall, 1980) and repetition (Howard & Franklin, 1988; Katz & Goodglass, 1990). Due to the restricted range of imageability ratings found in pictureable items, picture naming may be an inappropriate tool to use in investigating imageability in word finding. This problem notwithstanding, Nickels & Howard (1994) found a significant effect of imageability in four of the people with aphasia participating in their study with high imageability items being named better than low imageability items. It may be however that more of the participants in their study had an imageability effect and that this was not revealed due to reliance on a picture naming task. The usual pattern in aphasia is of an advantage for high imageability items over low.

There are few reported cases of a reverse imageability effect in aphasia. Marshall, Pring, Chiat & Robson (1996) report details of RG who had jargon aphasia following a left hemisphere stroke. On a number of tasks RG was significantly better at abstract than concrete words (word association task, word to picture matching, and naming to definition). Further investigations of RG's visual knowledge of items led these authors to conclude that RG had a deficit affecting visual properties within semantics.

There are a number of reported cases of people with progressive disease presenting with a reverse imageability effect. Warrington (1975) reported such an occurrence in the name definitions provided by one person AB with progressive deterioration of semantics (later described as having semantic dementia) who provided rich descriptions of abstract words but poor descriptions of concrete words. Breedin, Saffran & Coslett (1994) describe a reverse concreteness effect in DM who had semantic dementia. In providing definitions of abstract and concrete words normal people show an advantage for concrete. For high frequency items DM showed equivalent performance across high and low imageability sets. For low frequency items DM showed an advantage for abstract over concrete. A reverse concreteness effect was also found for comprehension and lexical decision.

Bird, Lambon Ralph, Patterson, & Hodges (2000) claim that the reverse concreteness effect found in the samples of connected speech from two people with semantic dementia are due to different distributions of frequency and imageability across nouns and verbs. Similarly Bird, Howard & Franklin (2001) ascribe an apparent advantage for content words over function words in reading aloud in three people with aphasia to an effect of imageability with high imageable content words being read aloud more successfully than low imageability function words. These findings form part of a reductionist move towards artefactual explanations of effects which were thought previously to demonstrate significant differential processing or storage along linguistic parameters.

1.4.1.3 Operativity

The term operativity was used by Gardner (1973) to differentiate items which are manipulable, and known to many senses (e.g. apple) from those which are known only to one sense (e.g. cloud). The notion of operativity lacks a clear definition however and

subsumes a number of potentially independent aspects of meaning including the various sensory modalities through which an object may be experienced and by which it may be represented cortically (see e.g. Coltheart, Inglis, Cupples et al's 1998 model of semantic memory incorporating multiple perceptual domains).

Few studies have investigated this aspect of meaning. Howard, Best, Bruce & Gatehouse (1995) looked at naming ability in 18 people with aphasia and found that independent aspects of the overall construct of operativity affected naming for some of the participants. Items which were deemed available to multiple senses or separable from the surrounding context were named better than those which were not. Manipulable objects however were named less well than non-manipulable objects.

1.4.2 Lexical variables

1.4.2.1 Frequency

Early studies of the effect of frequency on word retrieval were unequivocal. Oldfield & Wingfield (1965) and Wingfield (1967, 1968) reported a frequency effect in naming in normal participants. Butterworth et al (1984) and Howard et al (1984) reported a similar effect in aphasic naming. More recent reports incorporating analysis of additional variables such as imageability and length, have failed to find a consistent effect of frequency for all people with aphasia, demonstrating that this is not the ubiquitous symptom it was once thought to be and that different variables may be predictors for different individuals (Nickels & Howard, 1994; Nickels and Howard, 1995).

There are nevertheless a number of single case reports of people with aphasia who present with a frequency effect in word finding (EST: Kay and Ellis, 1987; EE: Howard, 1995; JS: Lambon Ralph, 1998; FR: Avila et al 2001). Kay & Ellis (1987) report a frequency effect which was present in spoken but not written naming.

There is one case report of a reverse frequency effect in aphasia. Marshall, Pring, Chiat & Robson (2001) reported details of JP's naming. JP was severely aphasic following a left hemisphere stroke. Her spoken output after onset was incomprehensible consisting of English jargon with neologisms. On a number of tasks JP showed a predilection for low frequency items shown in picture naming, naming to sentence completion, and

category fluency. In addition her semantic errors were lower in frequency than their target words (see section 1.3.1.7).

Although for many people with aphasia frequency may predict naming success, this is by no means a foregone conclusion. The locus at which the frequency effect is hypothesised to operate has been isolated to the lexeme or phonological word form level of spoken word production. Jescheniak & Levelt (1994) asked normal bilingual Dutch/English participants to translate words that produced either high frequency homophones (e.g. key) or low frequency homophones (e.g. quay). There was no difference in latency between high and low frequency homophones, leading the authors to claim that the low frequency homophones benefited from the accessing speed of the high frequency partner. As these share representations only at a phonological level, they suggest this level is implicated in frequency effects.

If a person's spoken word production problems do not arise at the lexical level one might suppose that they would be unaffected by frequency, thus reports of people with aphasia whose naming is not affected by frequency are not surprising.

1.4.2.2 Age of acquisition

A number of group studies of people with aphasia have found significant effects of age of acquisition on picture naming (e.g. Feyereisen, Van Der Borght, & Seron, 1988). In this study however the possible confounding effect of word frequency was not considered and the results are therefore equivocal.

Hirsh and Ellis (1994) published what they claim is the first single case study account of an age of acquisition effect in aphasia. NP had aphasia following a left hemisphere CVA. Analysis of spoken and written picture naming, reading aloud and repetition revealed an effect of age of acquisition in all tasks apart from repetition. Hirsh & Funnell (1995) subsequently published details of an age of acquisition effect in a case with progressive aphasia.

1.4.2.3 Familiarity

The familiarity of an item can be distinguished from the familiarity of the lexical term. It is important in using databases with values for familiarity that the method used to collect the data is known. In particular whether participants were asked to rate the word or the concept. Gernsbacher (1984) used the term ‘experiential familiarity’ to describe participants’ ratings of how familiar they were with written words. Nickels and Howard (1995: 1287) state that familiarity may be viewed as a subjective measure of word frequency.

In many reported case studies of aphasic naming familiarity is not controlled for. Nickels and Howard (1995) looked at the effects of eight variables on picture naming for two groups of people with aphasia. In the first part of the study they found a correlation between familiarity and aphasic naming success for eight out of 12 people with aphasia, but when logistic regression was carried out this was significant for only one person. As the authors point out, one reason for this is that familiarity correlates highly with word frequency and age of acquisition, and an apparent effect of one variable shown in a simple correlation may be caused by effects of other variables. Ellis, Lum and Lambon Ralph (1995) investigated the effects of a number of variables on naming success, but used concept rather than word familiarity. It is important to distinguish between these two in investigating the effect of familiarity.

1.4.3 Variables acting upon phonological encoding

1.4.3.1 Syllable and phoneme length

Word length effects are thought to operate after access to the abstract phonological representation has been achieved. In terms of the model shown in Figure 1.1 this implicates the phonological output buffer. Word length is measured in terms of the number of syllables or the number of phonemes. Many studies use these two variables interchangeably although as Nickels and Howard (2004) point out they implicate different processing routines in current models of speech production (e.g. Levelt et al, 1999). Models such as Levelt et al’s (1999) WEAVER ++ model (see Chapter Two) differentiate between processes acting on a syllabary of stored syllables, and subsequent processes involving phonological segments. In theory therefore breakdown might affect

either one of these and they are not homonymous processing routines. There are two distinct ways in which processing might be affected. Shattuck-Hufnagel (1979) described a slot and filler system of phonological encoding, where slots are made available for the relevant phonemes, which are then inserted into the slots in a left to right fashion. Disruption to such a system would mean slots are wrongly filled with misordered phonemes or are not filled at all, both resulting in phonological paraphasias. In contrast disruption to a short term memory store such as the phonological buffer shown in Figure 1.1 will result in part word productions with the word onset being better preserved than the offset and shorter words being produced more efficiently than longer words.

Nickels and Howard (2004) investigated the effects of number of syllables, number of phonemes, and syllabic complexity⁵, on word repetition for nine people with aphasia who had been identified as having a length effect in spoken naming. They used word repetition to assess spoken output. When they analysed syllable number independently they found people with aphasia were more accurate in production of two syllable than one syllable words, counter to the expected outcome. However, when they used logistic regression to identify the relative contributions of the three variables to the naming outcome they found only number of phonemes showed a consistent predictive contribution to naming outcome.

1.4.3.2 Group studies and single case studies

Effects on phonological encoding or later in production will result in phonological paraphasias. These may be real or non-words but will bear a phonological relationship to the target. Group studies have shown contrasting effects of word length. In an early investigation Goodglass, Kaplan, Weintraub and Ackerman (1976) found a relationship between success in naming and word length with decreasing performance with increasing length in a group of people with aphasia. Nickels and Howard (1995) found a length effect for only some of the people with aphasia they studied. A number of single case reports exist demonstrating length effects. The usual pattern is for longer words to be impaired relative to shorter words. Length effects are found often, although not exclusively, in the context of conduction aphasia. Franklin et al (2002) found a length

⁵ Syllabic complexity is the term used to refer to the number of clusters in the syllables of a word. Thus 'car' has a CV syllable whilst 'length' has a CVCC structure and thus is more complex.

effect in MB's spoken word production (see section 1.3.3.2). The relationship between phonological errors and word length is not straightforward. Gold and Kertesz (2001) report data from MMB who was aphasic, and who produced substantial numbers of phonological paraphasias in repetition. Analysis of word length and occurrence of errors showed no relationship between the two however.

There are very few examples of a reverse length effect in aphasic naming. Best (1995) reported the spoken word production of CJG who showed an advantage for longer over shorter words in picture naming only. This effect was not present in reading aloud or repetition. He made very few phonological errors in picture naming tending to fail to respond and then produce a semantic error.

1.4.4 Summary

In this section details of the variables thought to operate over people with aphasia's naming performance have been described. These fall into three broad categories: semantic, lexical, and phonological encoding. Studies of aphasic naming increasingly although not routinely report the influence of such variables over word finding. The methods used to analyse the impact have developed and now the results of multiple regression and logistic regression analyses are increasingly reported.

1.5 SUMMARY OF THE CHAPTER

In this chapter a brief overview of approaches to understanding aphasic word-finding from a classical aphasia syndrome account has been described, followed by a cognitive account of language processing and its breakdown in aphasia. The latter approach informs the methodology used in the work described in this thesis. Patterns of word finding have been described in line with the single word processing model shown in Figure 1.2. Methodological difficulties encountered in attempts to investigate this symptom of aphasia have been outlined. The impact of psycholinguistic variables on naming has been outlined. In the next chapter current influential models of spoken word production will be outlined and their ability to account for the findings from aphasia analysed.

CHAPTER TWO: MODELS OF SPOKEN WORD PRODUCTION

2.0 INTRODUCTION

In this chapter leading theories of spoken language production developed within the domain of cognitive science are outlined. Areas of overlap and debate within the set of theories are described. The adequacy of each theory's ability to account for the data from studies of aphasic speech production is discussed.

2.1 METHODOLOGY

The development of the metaphor of the mind as a representational system operating actively on incoming stimuli, and the notion of modularity, have combined to permit the development of a methodology for the study of mental function. The means by which cognition is studied will be described with particular reference to speech production.

2.1.1 Speech errors

Speech errors are a strong source of evidence for theory development in the area of speech production (e.g. Garrett, 1975, 1980; Dell, 1986; Dell, 1988). Proponents of this view believe that slips of the tongue can be examined to expose the otherwise hidden mechanisms of the system. This approach does have its limitations: a slip may be interpretable in two or more ways and prior theoretical assumptions may then dictate analysis decisions (Dell, 1995); perceptual characteristics of the auditory speech perception system may mean we only perceive certain types of error, for example content words.

An extension of the above methodology is that of experimentally induced speech errors first developed by Baars Motley & MacKay (1975). Subjects are given word pairs to produce under time constraints. This form of investigation allows the researcher closer control over the factors influencing error production and in particular has allowed quantification of certain phenomena such as the predominance of real word errors over non-word errors (so called lexical bias) (see e.g. Baars, et al, 1975; Dell & Reich, 1981).

2.1.2 Reaction time experiments

The analysis of speech errors is not deemed a valid method of investigation by certain researchers (e.g. Levelt et al, 1991a). Instead of looking at breakdown in processing an alternative approach is to measure reaction times, which reveal differential processing requirements and are generally used to illuminate aspects of temporal ordering in speech production. For example, Roelofs (1996) studied serial planning of morphemes, concluding from the time taken to produce primed morphemes that only the first morpheme can be planned ahead of production, subsequent morphemes having to await the processing of the first morpheme. The importance of this methodology to this area of research has only become technologically possible during the last few decades.

2.1.3 Cognitive neuropsychology

Further evidence regarding the nature of the processing subsystems comes from studying brain damaged performance. This area of enquiry assumes modularity and proceeds by investigating impaired and intact subsystems and deducing from these findings the nature of the normal system. Additionally, patterns of recovery of performance after brain damage contribute to this domain of inquiry. Models derived from neuropsychological findings (e.g. Ellis and Young, 1988) make no claim about the relationship between mind and brain.

2.1.4 Connectionism

In order to guard against bad theorising Flanagan (1996) argues for maximal consistency among levels of explanation, in particular between models of mental function and those of brain function. An area of cognitive science which has developed over the last 20 years and which attempts to unite these two is connectionism. The interconnecting neurone-like units articulated in computer programmes have similarities to the interconnections of the brain. The former are trained to equate a particular input pattern, for example a written word, with an output pattern, such as a string of phonemes (see e.g. Plaut & Shallice's 1993, implementation of word reading). Thus

normal processing can be modelled, and by ‘lesioning’ such a network⁶, impaired performance can also be investigated, thus shedding light on the nature of the impairment. An example of this is provided by Schwartz, Saffran, Bloch & Dell (1994) who simulated the speech of NC who had jargon dysphasia. Connectionist simulations of recovery of function after lesions (e.g. Martin, Saffran, Dell and Schwartz, 1994; Dell et al 1997: see Harley, 1996) are beginning to contribute to the understanding of normal cognitive functions.

2.2 THEORIES OF SPOKEN WORD PRODUCTION

2.2.1 Architecture: Boxes, arrows, units and interconnections

Theories of spoken word production are often represented in figures using box and arrow diagrams and owe their instantiation to analogies with computers: a set of processors operate upon the input signal and deliver an output which may have little surface relationship to the input signal. For example, picture naming involves visual input in the form of pictures, and verbal output. According to Morton (1985) the boxes are deemed to hold processors, and the arrows linking the boxes transfer output from one processor to the next processing module. Connectionist models differ from box and arrow models. Here sets of nodes or single processing units represent individual linguistic items. Nodes are connected within levels (e.g. semantic nodes in Plaut & Shallice’s (1993) model, and between levels (e.g. semantic units connect to word units in Dell et al’s (1997) account of spoken word production. Both types of model assume hierarchical levels of processing derived from psycho-linguistic theory.

2.2.2 Architecture: Semantic and phonological levels

Most models of speech production postulate at least two distinct levels of processing (Butterworth, 1980, 1989; Caramazza, 1997; Dell, 1986, 1988, 1989; Dell et al, 1997; Harley, 1995; Levelt, 1992; Levelt et al, 1999; Patterson & Shewell, 1987; Rapp and Goldrick, 2000). At one level a semantic representation of the item is accessed. This then accesses a phonological representation from which instructions for phonetic

⁶ Lesioning is achieved by one of several mechanisms: representational nodes are damaged or removed, connection weights are disturbed, decay rates of activation are increased or decreased, or noise is added.

realisation are derived. The evidence for the existence of these two levels comes from a variety of sources. Tip of the tongue evidence indicates that subjects in a TOT state have access to the meaning of the word, but cannot access the phonological form (Brown & McNeill, 1966). Speech error data reveal two broad types of errors occurring: whole word substitutions, (e.g. Fay & Cutler, 1977) and sound substitution errors (e.g. Garrett, 1975). These data have been taken as evidence that two distinct levels of processing exist. Evidence from aphasia provides further support for the existence of these two levels. Patients exist who show a semantic deficit in input tasks and confrontation naming (e.g. JCU reported by Howard & Orchard-Lisle, 1985) and others exist who make phonological errors in spoken output (e.g. EST reported by Kay & Ellis, 1987). Kempen & Huijbers (1983) labelled these two levels of representations 'lemma' and 'lexeme'.

2.2.3 Evidence for independent levels: semantics and phonology

2.2.3.1 Speech errors in normal speakers

The existence of two distinct types of speech error - whole word errors and sound errors - in naturally occurring and experimentally induced reports of speech errors, supports the independent representation of words at one level and sounds at a different level. These two types of errors differ in significant ways. In whole word exchanges such as "writing a mother to my letter" (cited in Dell, 1986) the exchanging units are typically phonologically distinct, share syntactic class, and may cross phrase boundaries. There are various types of sound error: exchange errors such as "The Lord is a shoving leopard" (an example from the Reverend Spooner, Eysenck & Keane, 1990), anticipatory errors, or perseveratory errors. The latter are unusual in normal speech production. Sound errors differ from whole word errors in that they typically involve only the first sound of the word involved, are from items which are proximate in the utterance and within the same phrase, and which do not necessarily share syntactic class.

Analysis of such errors led Garrett (1975, 1980) to propose distinct levels of processing within a model of spoken sentence production (see Figure 2.1). After activation from the message level the functional level generates a predicate argument structure along

syntactic constraints with slots ready for whole word entries to be inserted. Whole word

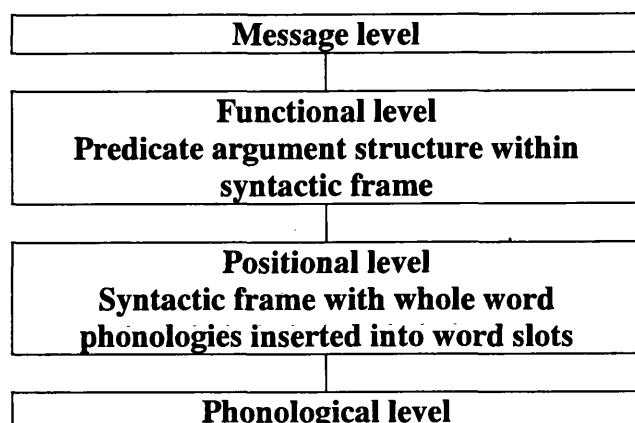


FIGURE 2.1 Garrett's (1975, 1980) model of spoken sentence production

errors occur due to mis-assignment of word entries to slots at this early stage in processing. The functional level in turn activates the positional level, where a syntactic frame is generated, into which the phonology of the already selected word entries is slotted. Whole word errors will not arise here as these are already assigned to slots from the previous stage of processing: sound errors may however occur, as the phonology of the selected word is retrieved.

Since Garrett's early work a number of researchers have investigated speech errors. In spontaneous speech they are surprisingly rare: roughly one per 1,000 words uttered (Garnham et al, 1981). There are difficulties however in coding these errors due to the fact that the intended target is not always transparent in conversational speech, and therefore the relationship between target and error is unquantifiable. Moreover listeners do not always attend to and notice errors, thus a sampling bias is suspected (Dell, 1995). In order to circumvent these problems researchers moved on to investigate speech error production experimentally. This is usually done by speeded repetition of word pairs. Dell & Reich (1981) used this paradigm to investigate the incidence of non-word errors. They compared subjects' errors in two types of word pairs: those which produced two real words when the first sounds were exchanged, and those which produced two non-words. They found that subjects were more likely to produce errors in the former type of pairs than the latter. This finding led Dell and others to an important claim regarding the nature of activation in the production system (see section 2.3.7.2). For now, suffice to say that speech error research has provided important evidence concerning the levels of

processing involved in spoken word production, and has opened a debate about the way in which the system controls error production.

2.2.3.2 Tip of the tongue states

Not only do normal speakers make errors in speech, they sometimes cannot produce the word they are trying to say. In a tip of the tongue state a speaker knows what she wants to say but cannot access the word form. Often speakers state that they know something about the word, in particular its first letter. They can also often describe the meaning of the word and provide synonyms of the target. Tip of the tongue states shed further light on the workings of the production system as they capture the system in a state of arrested operation.

Experimental studies of ToT are conducted by giving speakers a definition of a low frequency word. For example Meyer & Bock (1992) asked subjects to name words such as the following: “Something out of keeping with the time in which it exists”. Target: anachronism. Subjects are then asked to indicate when they are in a positive ToT state. They are then questioned as to the meaning of the word, asked to supply first, last and other letters, word length, and other words that are similar to the word in form. In a review of the area Brown (1991) summarised the findings from studies of the ToT state. Subjects report some knowledge of the word form: that is, processing is not completely arrested but rather information available is scant. Knowledge of the first letter is at around the 50% mark, which is above chance; knowledge of the last letter is also available, although other letter knowledge is limited. These studies lend further support to the claim that meaning is stored independently of form.

2.2.3.3 Reaction time experiments

As stated earlier the analysis of errors and ToT is not considered by some researchers to be a valid source of evidence in building a theory of how the system works, concentrating as such research does on system malfunctions. Levelt et al (1991) suggest that theories based on speech errors provide an account merely of how the system breaks down and not of how it works. The approach used by Levelt and colleagues involves measurement of processing times in reaction time experiments with non-brain

damaged participants, which reveal differential processing requirements within speech production.

Whilst most of this research illuminates the time course of the processing stages, the evidence also points to there being independent levels of representation. Levelt et al (1991a) asked normal participants to name pictures of items and, at different points after the presentation of the picture, words were presented auditorily for lexical decision. The words were related to the picture target either phonologically or semantically. For example a picture of a sheep would in one condition be accompanied by the word “goat” and in the other condition by the word “sleep”. These words were presented at different stimulus onset asynchronies. Measurement of latencies for the lexical decision task led to the conclusion that there is early but not late activation of semantic neighbours of target words: lexical decision at late stages post presentation of the picture was not speeded suggesting the target was not activated. Measurement of latencies showed no advantage to reaction times for phonological neighbours such as “sleep” when these were presented early on, but there was an advantage when these appeared late. These data add further support to the notion that independent temporally ordered processing of semantics and phonology occurs.

2.2.3.4 Evidence from aphasia

The two cases mentioned in Chapter One (JCU, Howard & Orchard-Lisle, 1984; EST, Kay & Ellis, 1987) are often cited as cognitive neuropsychological evidence in favour of independent levels of semantic and phonological representations (e.g. Lesser, 1989; Ellis, Franklin & Crerar, 1994). Since those two people with aphasia were described many more people with a semantic deficit have been reported (e.g. KE described by Hillis Rapp Romani & Caramazza, 1990) and with a phonological deficit (e.g. MF described by Best, 1996).

On closer examination however these cases are not so clear cut, for example EST also made semantic errors in picture naming (see Nickels, 1997, for a reanalysis of the evidence). Moreover it is not the case that people with aphasia routinely present with discretely impaired processing attributable categorically to a breakdown at one level of processing: many people with aphasia present with impaired processing at a number of

levels. Nevertheless the clarity of the distinction supported by the evidence from normal speakers remains.

2.2.4 One-step models

Recent models of production (e.g. Caramazza, 1997; Dell et al, 1997; Levelt et al, 1999; Rapp & Goldrick, 2000) agree that in turning meaning into sound a semantic representation is addressed which accesses a corresponding phonological form. The single step involved in this translation has led to models which incorporate only these two levels of representation being termed 'one-step models'.

This one-step model has dominated research and clinical work in aphasia. Many researchers believe however that this model is under-specified, thus the assessment and treatment currently being offered rely upon an inadequate description of the problem. People with aphasia exist who appear, on the evidence of assessment based on this theory, to present with the same locus of impairment in function, but who show different patterns in error production and in response to treatment (see e.g. Best et al, 2001; Hickin et al, 2002). Thus it could be argued that clinical provision to people with aphasia relies upon an inadequate theoretical model.

2.2.5 Two-step models

2.2.5.1 Lemma, or word level

Whilst there is broad agreement concerning the existence of independent semantic and phonological layers of representation, disagreement exists regarding the existence or otherwise of an intermediate level of representation, known as the lemma level. Kempen and Huijbers (1983) first proposed the lemma as a semantically and syntactically specified representation mediating between conceptual semantics and phonological representations. According to Levelt et al (1999) however the lemma mediates between lexical semantic representations and the phonological level, and at the lemma level information about an item's syntax is made available. Thus, in the resulting two-stage model of lexicalisation, access to a word's phonology, or lexeme, is dependent upon accessing the syntactic representation or lemma.

Serious debate surrounds this issue, and current studies have focused on the independence of syntactic information (lemma level information) from phonological or word form information (lexeme level information). If a lemma level exists, access to phonology will be dependent upon access to this prior stage. Therefore syntax will be available before phonology, and in ToT states syntactic knowledge will be available when phonological or word form knowledge is not. Evidence exists to support the contention that the two forms of information are independent. The issue of whether serial access occurs remains unclear.

2.2.5.2 The independence of syntax (lemma) and phonology: tip of the tongue studies

Evidence in support of the independence of syntax and phonology comes from studies of normal subjects in tip of the tongue states (Vigliocco, Antonini and Garrett, 1997; Caramazza and Miozzo, 1997). Here subjects are given definitions of low frequency words which elicit a tip of the tongue state. They are then asked questions about items they are unable to name, which probe knowledge of the word's syntactic properties (typically grammatical gender, but more recently phrasal structure properties such as those governing the use of count versus mass nouns in English) and knowledge of the word's phonological properties (first and last sound, number of syllables). These studies have shown that subjects in a ToT state are frequently able to report the grammatical gender of items they cannot name, knowledge of the words' phonological form being available partially or not at all (Vigliocco et al, 1997; Caramazza and Miozzo, 1997). Similar studies have been reported probing the count-mass distinction in English nouns. Vigliocco, Vinson, Martin, and Garrett (1999) report that normal subjects have knowledge of the count-mass status of items for which they are in a ToT state.

These results have been interpreted as support for the existence of an intermediate stage of processing at which syntactic information is available. Lexically represented syntactic features are available when the full word form is not. These findings support the notion of independent representations for syntax and phonology. They do not determine whether serial access to syntax then phonology occurs (as the two stage model holds), and consequently whether access to syntax is obligatory in the process of lexicalisation.

2.2.5.3 The independence of syntax (lemma) and phonology: Evidence from neuropsychological studies

Further evidence to support the independence of syntax and phonology comes from neuropsychological studies. Henaff-Gonon et al (1989) report a single case study of a French anomic subject, GM. GM presented with fluent speech with marked word finding problems. His responses often consisted of commentaries around the semantic field of the target, and sometimes incorporated details about the word's gender, phonology and written form. Of interest here is that he was able to correctly identify the grammatical gender of items whose name he could not produce (13/14 test items). Thus, syntactic information was available when the phonological form was not. (Systematic testing of his access to phonological information was not carried out although phonological information was sometimes available).

Badecker, Miozzo and Zanuttini (1995) report a further single case known as Dante. Dante presented with anomia as a result of a meningo-encephalitis. Dante was presented with pictures he had been unable to name in a picture naming experiment, and was then asked to identify the grammatical gender of the item, and to indicate the word's first and last letter, its length, and any other word that was phonologically or orthographically similar. Dante was able to identify the gender for 106/111 items, but was unable to give any information at all about the phonological form – his inability to complete this task was so profound that the task was abandoned. The authors interpret these results as evidence in support of the two-stage model: Dante has access to the lemma level, but access to phonology is impeded. Further information about Dante's access to syntactic knowledge but inability to access phonology is provided in a paper by Miozzo and Caramazza (1997), in which they describe his ability to identify the auxiliary of verbs he was unable to name.

More recent evidence from a single aphasic subject also supports the hypothesis that syntactic and phonological information are separately represented. Vigliocco et al (1999) describe the performance of MS who was able to identify whether an item was a count or mass noun even though he was unable to name the item in a picture naming task. Additionally MS was able to access some phonological information, although this did not appear to be dependent upon prior access to syntax.

All three cases support the contention that syntax and phonology are independent and that they are represented separately. The studies are equivocal however on the issue of whether serial processing of syntax then phonology occurs (although Dante's results were interpreted in this way), and thus whether access to phonology is dependent upon access to syntax. The evidence from Dante comes closest to providing clear evidence in this respect: Dante was able to access syntactic information but no phonological information. Caramazza (1997) proposes that the reports outlined above demonstrate the independence of syntax from phonology but tell us nothing about the relationship between the two.

2.2.5.4 The organisation of syntactic and phonological knowledge

The above data provide evidence that syntax and phonology are independent of each other but do not shed any light on their organisation. One set of studies exists which directly addresses the issue of serial organisation of syntax and phonology. Van Turennout Hagoort and Brown (1997, 1998) carried out ERP studies in which participants were asked to make decisions based on gender and phonological information in a picture naming task. Results were clear: participants were able to access a syntactic property (gender information) without demonstrating any evidence that they were preparing a response based on phonological information. That is, processing could be arrested after access to syntax. The reverse was not found: participants were not able to prevent gender information being activated in the case of making phonological judgements. It remains the case that these data provide the strongest evidence of serial organization of syntax and phonology and that the data from this unusual research paradigm need to be backed up by converging evidence from other sources.

2.2.6 Activation within one and two step models

Activation can be feed-forward or interactive. In the former activation moves forward only through the system. In the latter activation feeds forward and backwards through the system. In models with discrete processing, processing at one level is complete before activation of the next level begins. In such theories (e.g. Levelt et al, 1999) there is no temporal overlap between stages. In models with cascading activation processing

overlaps in time: processing at one level can commence before it is completed at the previous level.

Discrete two-stage models of production (e.g. Butterworth, 1989, 1992; Levelt, 1992; Levelt et al, 1999) assume completion of processing at one level before processing begins at the next. Morton's (1980) logogen model derived from normal and neuropsychological data is an example of a one step model incorporating discrete processing. Dell's (1986) model and subsequent versions of the same model (Dell, 1986; 1988; 1989; Dell et al, 1997) differ in assuming temporal overlap in processing between adjacent stages.

Rapp & Goldrick (2000) propose substantial modifications to the complete interactive activation account proposed by Dell and colleagues. They propose no feedback from the lexical level to the word level, and weak feedback from the phonological to the word level within a model that is otherwise the same as Dell et al's (1997) DSMSG model.

2.2.7 Four models of spoken word production

2.2.7.1 *WEAVER ++ (Levelt, Roelofs and Meyer, 1999)*

Levelt and colleagues have been refining this theory for a number of years. The 1999 version is the latest attempt to provide an account of normal spoken word production. The architecture of the model is shown in Figure 2.2.

The model involves feed-forward discrete processing. It incorporates spreading activation within layers, which then flows to the next layer. Processing at one level is completed before it begins at the next. The conceptual stratum consists of nodes some of which represent lexical concepts. These are non-decompositional units of representation.

Importantly, although representations within the conceptual stratum do not share sets of features, an activated lexical concept will nevertheless activate another semantically related concept through connections between nodes within the layer. Active lexical concepts (i.e. that of the intended target and those of semantically related neighbours) will spread activation to their corresponding lemma nodes. The highest activated lemma

will be selected, and under most conditions its syntactic information made available⁷. This enables the appropriate syntactic framework for the word to be created. In the case of nouns the syntactic information will include the fact that it is a noun, whether it is

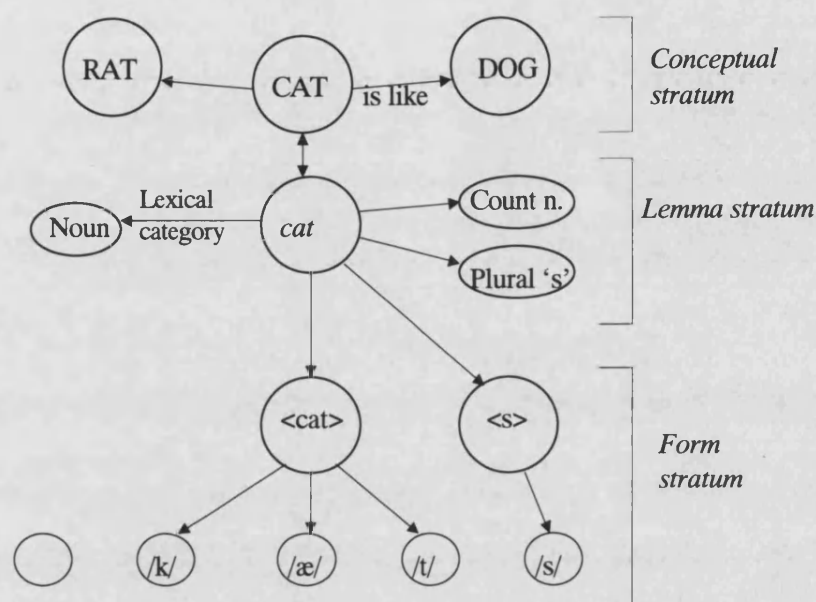


FIGURE 2.2 WEAVAR ++ adapted from Levelt, Roelofs and Meyer (1999)

count or mass, how the plural form is computed, and, in languages which mark grammatical gender, the gender of the noun. The selected lemma now activates the form stratum. In the theory this proceeds in three levels: morphological, metrical shape and segmental makeup. Only one lemma is selected and activation of the corresponding lexeme proceeds directly from the selected lemma. This is in line with Levelt et al's (1991a) findings of no phonological activation of semantic neighbours of the target (see section 2.3.7.2 below).

Evidence for the theory comes entirely from chronometric experiments with normal participants, where the data consist of production latencies under particular experimental conditions. These authors do not consider speech error data, tip of the

⁷ A gender priming effect is only found when participants are asked to produce noun phrases with an adjective (Schriefers, 1993). When asked to produce bare nouns there is no priming from the distractor word's gender (Jescheniak, 1994). This led to the conclusion that in certain circumstances syntax is not activated (see Levelt et al, 1999, for a discussion).

tongue data, and aphasic data as valid evidence for developing theory in this domain. They claim that building a theory of how something works efficiently cannot be done by using evidence from how the system goes wrong (although they do state that “the theory should be able to account for error patterns as well as for production latencies”: Levelt et al, 1999: 3).

2.2.7.2 Dell, Schwartz et al (1997): *Interactive two-step model*

This model includes a semantic level, a word or lemma level (the two terms are used interchangeably) and a phonological level (see Figure 2.3). The semantic level consists of decomposed representations where a concept is represented by a pattern of activation across a set of semantic features. Concepts similar in meaning will share semantic features. When a concept is activated that activation will spread forward through the model to the word or lemma node. These connections from semantic features to the lemma are excitatory and bi-directional. Like Levelt et al’s account the lemma is a unitary node. Unlike Levelt et al’s account, the lemma is a semantic *and* syntactic representation. Phonology is segmental, relevant segments being activated from the active lemma node, again through excitatory connections.

Crucially the connections are bi-directional and activation thus flows forward through the model, *and* backwards. Thus it is an interactive activation model. This means that activation spreads from semantic units to word units and to phonological units, and from word units to semantic units and from phonological units back to word units. Feedback in the network means that the activated word node spreads activation back to the semantic level thus activating neighbours of the intended target, and the activated phonological segments spread activation back to word nodes activating phonological neighbours of the target. In the original model (Dell, 1986) the spread of activation between layers was unbounded. In a revision of the model Dell and O’Sheaghda (1991) proposed limited activation to adjacent layers.

Interactive activation was proposed by Dell (1986) in the original model to explain two phenomena found in naturally occurring errors and subsequently in experimentally induced errors. These are the mixed error effect and lexical bias (see also 2.3.3.1). The former refers to the finding that speech errors occur which have both a semantic and a phonological relationship with the target at a rate that is higher than would be expected

by chance. The latter refers to the finding that speech errors are words rather than non-words again at a rate that is higher than would be expected by chance. Several studies have shown evidence of mixed errors.

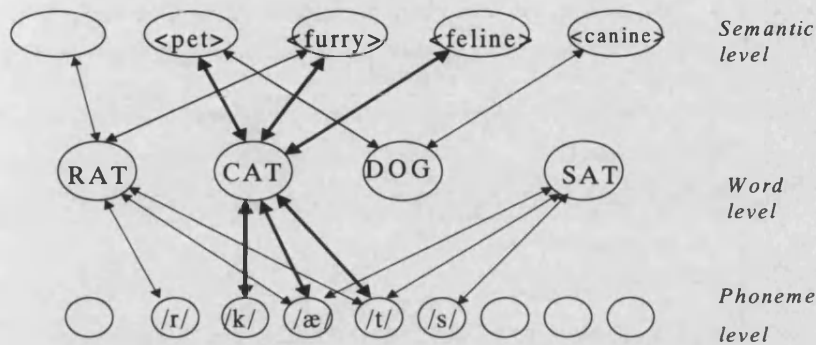


FIGURE 2.3 Interactive activation model of Dell, Schwartz et al (1997)

The model explains this as follows. In producing a word such as oyster forward activation from the semantic to the word level activates all items sharing semantic features with the target word. Thus lobster, crab and so on will receive some activation, along with lobster. Feedback from the word level to the semantic level will activate semantic features of the target, plus further semantic features relating to the semantic neighbours (such as lobster and crab). Feed-forward activation then activates the target and related items at the word level. Feed-forward activation from the word level to the phoneme level will activate segmental information relating to the target and semantic neighbours. Feedback from phoneme to word level will activate the target word node plus nodes of words sharing phonological segments with the target. The summation of this activation from both forms of interactive activation will lead to items which have both a semantic *and* a phonological relationship to the target being selected over other activated items. Thus crab shares just a semantic relationship with oyster, and bolster shares just a phonological relationship, whereas lobster shares both. A similar explanation holds for lexical bias: phoneme errors resulting in non words, such as blue bug -> blue blug (Dell, 1986) will receive no activation from the word level. Phoneme errors resulting in words will however be activated from the word level as feedback from the phoneme level will activate the word level in turn reactivating the phoneme

level. As for mixed errors the word string will be selected as it will be higher in activation.

The rationale for the inclusion of the lemma or word level comes from two sources. First, within connectionist modelling an important difference exists between regular and non-regular mappings of representations in adjacent layers. In the former, where two items share characteristics within one level of representation - e.g. the graphemes of two written words such as cat and can - they will also share a similar number of characteristics in the next layer of representation, e.g. phonology. Thus cat and can share two graphemes, and also two phonemes. In this case the mapping can be achieved directly without intervening layers. Where mappings are not regular however the sets of features for two similar items in one layer or level of representation will not map onto two sets of similar features in a second layer, e.g. dog and cat share many semantic features but no phonological features. The mapping is thus arbitrary. In a connectionist model such mappings necessitate the incorporation of a hidden layer, which in the case of spoken word production may be served by the lemma level (Dell et al, 1997).

The second reason concerns the model's ability to explain the syntactic class constraint found in the overwhelming majority of form related errors (Fay and Cutler, 1977). These errors are similar to the target in form, but also share syntactic class. If they arose due to encoding problems at the phonological level there would be no syntactic influence over the outcome error. The influence of the lemma level however, through activation spreading back to that level and then forward again to the phonological level, constrains the error to one of the same syntactic class.

2.2.7.3 Rapp and Goldrick's Restricted interactive activation (RIA) model

Rapp and Goldrick (2000) considered five possible accounts of spoken word production and the ability of each to account for four types of data. These were: the occurrence of mixed errors in non-aphasic speech; lexical bias in the speech errors of non-aphasic speakers; the occurrence of only semantic errors in two aphasic speakers, one with and one without a semantic deficit; and two different patterns of phonological relatedness between target and semantic error. Details of the latter set of data are as follows: where semantic errors arise within semantics, as is the case for aphasic participant KE, no evidence of a phonological relationship between target and error was found; where

semantic errors arise after access to semantics is complete i.e. within or in access to a lexical or lemma level, as is the case for aphasic participants PW and CSS, a phonological relationship between target and semantic error was found.

The Restricted Interactive Activation model (RIA) proposed by Rapp and Goldrick (2000) is identical to Dell et al's (1997) model in terms of levels of representation (see Figure 2.3). It differs however from this model in terms of activation. There is no feedback between the lexical and semantic levels, and only restricted interaction between the phonological and lexical levels. Using computer simulations Rapp and Goldrick (2000) compared each model's ability to account for the findings from aphasic and non-aphasic data and concluded that the RIA model provided the best account of the data. A discrete feed-forward account such as that proposed by Levelt et al (1999) fails to account for mixed errors and lexical bias in normal speakers, and for phonological relationships between target and semantic error in certain patterns of aphasic production. Some interaction between the lexical or lemma level and the phonological level is required. The highly interactive account of Dell and colleagues (Dell, 1986, 1988; Dell et al, 1997) on the other hand could account for all of the data but only by restricting interactivity, particularly at the upper level between the lexical level and the semantic level. They conclude that interaction at this level is "at best irrelevant and at worst a liability" (Rapp and Goldrick, 2000: 486).

2.2.7.4 Caramazza's (1997) Independent Network model

This model (Figure 2.4) differs from the three outlined above in dispensing with an obligatory intermediate level between meaning and sound, but includes written output as well as spoken. There are three levels of representation: a lexical semantic network, a syntactic network, and phonological (and orthographic) lexemes⁸. Access to the syntactic network is not obligatory in translating from meaning to sound (or spelling). The lexical semantic network consists of semantic nodes. Activation flows forward to the corresponding phonological lexeme, which then activates the necessary segmental information for that word. Syntactic information can be accessed but this is not necessary in order to achieve the mapping from semantics to lexeme. Having argued

⁸ The model attempts to account for the data from aphasic participants. It includes spoken and written output. For simplicity's sake the discussion here is limited to spoken output.

forcibly against the necessity of the lemma, Caramazza somewhat confusingly states that the lexemes are “semantically and syntactically specified lexical representations” (Caramazza, 1997: 196). It is hard to see here how these differ from Dell et al’s (1997) word/lemma nodes.

Within layers of the model there are inhibitory connections. In producing the Italian word “tigre” semantic nodes activate the relevant lexeme which may under certain conditions activate the syntactic network nodes of noun, feminine, and count. Other grammatical class nodes such as verb and adjective, the masculine gender node, and the mass noun node will all be inhibited within the network. The model was constructed to account for the data from aphasic participants, in particular the occurrence of grammatical class deficits in one output modality only (Caramazza and Hillis, 1991; Hillis and Caramazza, 1995; Rapp and Caramazza, 1997; Rapp, Benzing and

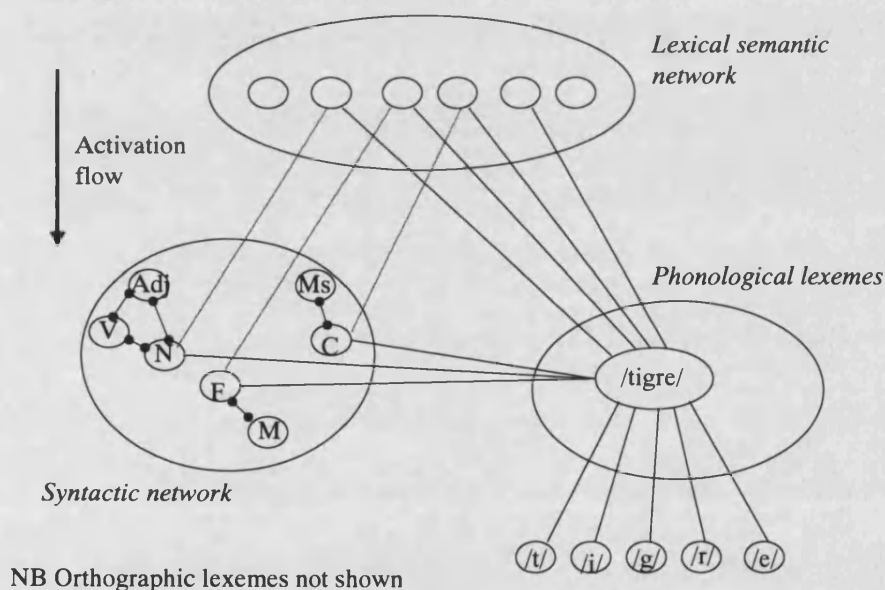


FIGURE 2.4 Caramazza’s (1997) Independent Network model

Phonological lexemes connect to the syntactic network. Words are coded here for grammatical class: Adj = adjective; N = noun, V = verb; for grammatical gender: F = feminine, M = masculine; and for noun status: C = count noun; Ms = mass noun.

Caramazza, 1995). The occurrence of grammatical class deficits, typically shown where a person with aphasia shows an advantage for naming nouns over verbs or vice versa, indicates that one of the organising principles of the lexical system is syntactic. A deficit in one modality only (either spoken or written) cannot be accounted for by a

modality-neutral lemma (proposed by Dell et al, 1997; and Levelt et al, 1999) as, if the lemma layer were damaged the grammatical class deficit would be evident in both output modalities.

2.3 LEVELS OF REPRESENTATION

2.3.1 Semantics

In accordance with the evidence from speech errors and from tip of the tongue research, theoretical accounts of the processes involved in translating meaning into sounds incorporate a semantic representation. There is disagreement about the architecture of the system, in particular whether semantic representations are decomposed, consisting of distributed patterns of activation across sets of semantic nodes, or non-decomposed consisting of lexical-conceptual nodes.

2.3.1.1 *Two systems or one?*

Dell et al (1997)⁹ propose a single semantic system which maps onto the word level. In this model semantics are not explicitly lexical. Levelt et al (1999) on the other hand propose a specifically lexical conceptual system which forms a sub-section of the larger conceptual system. Caramazza's (1997) model has a lexical conceptual system which is similar to that of Levelt et al (1999).

In a non-specific semantic system, which incorporates non-lexical conceptual information and lexical information, the preservation of non-lexical semantics along with impaired lexical semantics (seen in many people with aphasia, such as JCU, Howard and Orchard-Lisle (1987) are difficult to explain. For these people object concepts are unimpaired but lexical concepts are damaged, as shown by their impaired performance on word to picture matching and other tests of input. If conceptual semantics are damaged they should show impaired object matching, sorting and use, such as is seen in people with more general semantic loss suffering from Alzheimers's disease or semantic dementia.

⁹ As Rapp and Goldrick's model is based on Dell et al's (1997) account it will not be referred to directly.

2.3.1.2 Semantic representations

The second issue concerning semantics relates to the nature of the representations themselves. In Dell et al's (1997) and Caramazza's (1997) accounts items have

TABLE 2.1 Semantic representations in three models of spoken word production

Authors	Lexical or non-specific	Decomposed or non-decomposed	Stages of processing between meaning and form
Caramazza (1997)	Lexical	Decomposed	One
Dell et al (1997)	Non-specific	Decomposed	Two
Levelt et al (1999)	Lexical	Non-decomposed	Two

distributed representations, and thus two semantically similar items share many semantic feature nodes. Levelt et al (1999) propose non-decomposed representations, thus semantically similar items are linked via activation within the semantic or conceptual layer (see Figure 2.1). Details of the three models' semantic architectures are provided in Table 2.1.

2.3.2 The lemma or word level

There is some debate as to the nature of the representations at this level. According to Levelt et al (1999) the lemma level allows access to syntactic nodes but that access is not obligatory in processing single words. The syntactic nodes consist of the word class of the target (noun, verb etc.), and in the case of a noun whether it is mass or count, the plural form, and in some languages the grammatical gender. Dell et al (1997) propose that the word level in their model also incorporates syntactic information. Caramazza differs from these two in proposing that his phonological lexeme is modality specific and does not necessarily activate syntactic information, although it may do so if required.

2.3.3 Morphological representations

In most models there is no mention of morphology. Levelt et al (1999) incorporate a level within the phonological stratum however, which they term morphology. The lemma representation maps directly onto this level which appears to contain nodes representing stems and affixes. Little information is given about these representations and why they form part of phonology rather than syntax.

2.4 PHONOLOGICAL ENCODING

2.4.1 Slot and filler model

According to Levelt (1992) and others the second stage of lexicalisation consists of phonological encoding. The main model describing processing at this level is the slot and filler model first proposed by Shattuck (1975) and developed in later work (Shattuck Hufnagel, 1979, 1987). Evidence for this model comes again from speech error data. Most sound errors involve word onset consonants (Stemberger, 1983) which slip to a similar position in the interacting unit (MacKay, 1972; Shattuck-Hufnagel, 1980; Stemberger, 1983). Shattuck-Hufnagel (1979) interpreted these findings as evidence that phonemes are allocated to slots in a frame, rather than words being encoded as a whole unit. Such errors, known as movement errors, provide strong evidence for this structure-content distinction. According to Dell (1995) and Levelt (1992) words are constructed time and again, the eventual form of the 'phonological word' being dictated by the surrounding context. In contrast, Dell, Juliano and Govindjee (1993) propose a parallel distributed processing model of phonological encoding. This model produces phonologically constrained errors but no movement errors, and therefore fails to account for all the data.

Speech error data provides evidence for the role of the word and the morpheme as planning units. Errors involving both these units are commonly reported. In contrast few errors involving syllables have been reported and, according to Shattuck-Hufnagel (1979), there is insufficient evidence to support the idea of the syllable as a processing unit. However, when sound errors occur they are constrained by syllable position, most notably, onset consonants shifting to onset position. This is known as the syllable paradox and poses a challenge to theorists. Shattuck-Hufnagel (1987) proposed two

hypotheses governing phonological encoding: representational units which correspond to single phonemic segments, and a mechanism governing the serial ordering of the segments. According to most theories, the syllable is proposed as the frame into which phonemes are slotted (e.g. Dell, 1986).

2.4.2 Serial ordering in phonological encoding

Evidence for the order in which morphemes and syllables are pre-planned and the order in which phonemes are allocated to their slots in the frame comes from reaction time experiments.

Meyer (1990) investigated phonological encoding in disyllabic and polysyllabic words by means of an implicit priming paradigm. Subjects learnt pairs of Dutch words, a link and a target word. The link word was then presented to the subject who produced the target in the pair as quickly as they could. The target words were selected to investigate the effect of particular similarities, for example a shared first syllable (kabel, kater), a shared second syllable (salto, veto). Facilitated production of the target in any condition is considered as evidence that the subject has been able to pre-plan that particular item. In Meyer's experiments, facilitated production occurred for the words sharing the first syllable only (kater, kabel etc) and not for words sharing second syllables. Meyer claims from these results that pre-planning of syllables occurs in left to right fashion incrementally: only when processing of the first is complete can processing of the second commence.

2.4.3 Temporal aspects of phonological encoding

The time course of phonological assignment to slots in the planning frame has been investigated by reaction time experimentation. Wheeldon and Levelt (1994) plotted the time course of successive phonemes in bisyllabic Dutch words (e.g. rustig). Their findings suggest that phonemes are activated in linear fashion from left to right, proposing an "unfolding phonological representation of the target word". These results support the view of Meyer and Schriefers (1991). Shattuck-Hufnagel (1987) had earlier proposed differential processing of onset consonants, with these being allocated to their position after everything else. The results of Wheeldon and Levelt's (1994) experiments appear to contradict this assertion. In Dell's (1986) original model phonemes were not

encoded in linear order but were marked for syllable position, and assigned to their positions in any order. Wheeldon and Levelt (1994) consider that minor adjustments to Dell's model are necessary to make their data compatible with an interactive activation account of phonological encoding.

2.5 MODELS OF SPOKEN WORD PRODUCTION AND APHASIC NAMING ERRORS

2.5.1 Background

The most stringent method of assessing a model's ability to account for particular patterns of breakdown whether from normal or aphasic speakers, is through a computerised implementation of the model. This is trained to produce the desired output and then lesioned to approximate the pattern seen in the aphasic speech (Dell et al, 1997; Foygel & Dell, 2000; Rumel & Caramazza, 2000). Such simulations are still in their infancy and at this stage are only able to account for error patterns in output without relating these to hypothetical levels of breakdown in overall processing. The success of such simulations in producing the aphasic error patterns are described in section 2.6.3 below. Where a computerised simulation has not been implemented one is left to intuit how a model might explain the data from aphasic speakers. This is addressed in section 2.6.2.

2.5.2 Non-implemented explanations

2.5.2.1 Semantic errors in input and output

The two stage model (Levelt et al 1999) has little trouble accounting for cases where semantic errors are produced in input and output tasks¹⁰. A deficit at the conceptual stratum accounts for both sets of symptoms. A semantic error is produced when the incorrect lexical semantic representation is selected, both for input and output tasks.

¹⁰ Levelt et al (1999) have provided data on the computerized simulation of WEAVER++ but have not lesioned this model to mimic the data patterns found in aphasic speech, thus this is treated here as a non-implemented account

In Dell et al's (1997) model semantic errors arise in the model because of the interaction between the semantic level and the word level: feedback from the word level to the semantic level activates semantic neighbours of the target who share semantic nodes with the target. Furthermore, errors are constrained by syntactic class because of the influence of the word level on selection. Evidence for this last claim comes from studies of aphasics' word selection errors which tend to obey a syntactic class constraint (Gagnon, Schwartz, Martin, et al 1997). Damage to the conceptual level would lead to semantic errors in input and output. Damage to the semantic-word interactive system accounts for output semantic errors. For damage here to also cause input semantic errors demands the assumption that the route between word and semantics is shared for input and output. In Caramazza's (1997) model damage to the lexical semantic network would engender both input and output semantic errors.

2.5.2.2 Output semantic errors without input semantic errors

RGB and HW described by Caramazza and Hillis (1990: see section 1.3.1.5) made semantic errors in picture naming but were able to carry out input tests of semantics without error. The authors explain this in terms of a response blocking mechanism: the phonological entry for the target is unavailable so the nearest item is accessed in its place. Thus semantic errors are deemed to arise post semantically in a model positing direct activation of phonological output representations from semantics (see Caramazza, 1997). The errors cannot be due to lexical semantic level damage as input tests are completed successfully. In accessing the phonological output representation from the lexical-semantic representation semantic errors arise. As access to syntactic information is not obligatory in this model the prediction is that word class will not constrain the production of errors.

As Levelt et al's (1999) and Dell et al's (1997) models propose a word or lemma level in between the lexical semantic level and the phonological level it seems logical to suppose that the pattern of deficits and intact performance seen in RGB and HW arises, according to these theories, at the lemma level or in accessing the lemma level¹¹.

¹¹ It is not feasible that semantic errors occur after access to a lemma has been achieved, in the route between lemma and lexeme. In the serial stage model of Levelt et al (1999) the discrete feed-forward mechanism involves completion of processing at one level before processing at subsequent levels occurs. Thus semantic effects after the lemma has been accessed are not possible.

Consequently in aphasic speakers like RGB and HW, who present with semantic errors in output only, certain other impairments should be present. Notably, if the lemma level is damaged or access to it is disrupted these aphasic participants should show impaired knowledge of syntax for items they are unable to name. In English this would amount to poor knowledge of plural, or misuse of plural forms, and impaired knowledge of phrasal structure rules (such as the count-mass distinction and syntactic rules relating to this; verb noun agreement for singular versus plural nouns; referential pronoun agreement for singular versus plural nouns).

2.5.2.3 Evidence for lemma level deficit

Of the neuropsychological studies which have investigated participants' knowledge of syntactic elements, none of the cases had a lexical semantic deficit and none appear to have made semantic errors. GM (Henaff-Gonon et al, 1989) rarely produced a semantic paraphasia either in connected speech or in picture naming, and comprehension of single pictureable items was apparently intact. Dante (Badecker et al, 1995) presented with intact semantics as shown by his performance on input tests such as word to picture matching and although the error types are not reported in full the strong implication is that failures to respond predominated. For MS (Vigliocco et al, 1999) semantics are assumed to be intact as shown by his normal comprehension on aphasia tests, and his speech errors were mainly substantial circumlocutions, with no report of semantic errors. Thus for none of these cases is there any reason to locate the source of their deficit at or before the lemma level. In all cases access to syntax is intact and the deficit lies in accessing phonology thereafter.

2.5.2.4 Form related errors

The occurrence of form related errors in aphasic speech is not difficult for all three theories to handle. All propose an independent level of phonological segments: after the lemma (Levelt et al, 1999) the word (Dell et al, 1997) or the phonological lexeme (Caramazza, 1997) has been activated the appropriate segments are inserted into the phonological frame. Only Levelt et al (1999) are explicit about this stage, postulating that the metrical framework for a given word will vary according to its sentential context, and that what is generated is a phonological word within which a given target word will be slotted.

The preponderance of real word errors in both non-aphasic speech and aphasic speech is dealt with satisfactorily by Dell et al (1997) as discussed above: interaction between the word and phoneme levels results in real words being more activated than a non-word phoneme string which will receive no activation from the word level. Neither of the other models account for this phenomenon explicitly.

2.5.3 Computerised simulations of aphasic naming

2.5.3.1 Dell et al (1997)

Dell et al (1997)'s simulation of spoken word production uses an interactive two step model. The lesioned model simulates the naming and repetition performance of a group of 21 people with fluent aphasia. The effects of re-learning on the lesioned model are also examined in order to model recovery. One assumption in this work is the continuity hypothesis (Freud, 1958) which proposes that normal speech and aphasic speech occupy positions on the same continuum. This assumption allows modelling of aphasic production to be legitimately reproduced by a model of normal production which is then lesioned. As Dell et al (1997: 811) report "If the continuity thesis is correct, the model should characterise these deficits (aphasic) without a great deal of added complexity".

To set up the model the authors trained it on two neighbourhoods with six words in each. All words had CVC structure. Each neighbourhood contained a target plus semantically related, formally related, and unrelated words, and in one neighbourhood a semantically and formally related word. The authors identified estimates of the error opportunities available to each error type, based on what a random response to a selected target would be. Taking the total responses they then were able to calculate the proportion of each error type for a target word. The model was then set up to mirror this pattern of performance.

Data from normal speakers' performance on the Philadelphia Naming Test (PNT: Roach, Schwartz, Martin, et al, 1996) were used to parameterise the model to fit control data. The normal controls produced semantic, formal, non-word, mixed and unrelated errors, and a small number of no responses and circumlocutory responses. Semantic, mixed, and unrelated errors were the most frequent types. Dell et al (1997) claim that the preponderance of whole word errors in the normal control data indicates a lemma

level source of errors in non-aphasic speakers. They propose that in lemma access the target's competitors are semantically activated due to interactive activation between the semantic and word levels. They claim that the influence of the phoneme level on lemma selection is not strong enough to override semantic activation, thus purely phonologically related errors are rare. However, the claim that interactive activation occurs between the word and phoneme levels is supported by the large number of mixed errors found in the normal control data.

The model was then fitted to the aphasic naming data. Two forms of damage to the model were applied: damage to connection weights, giving decreased coherence between levels; and increase to the decay rate, giving a loss of integrity to representations. The model assumes global damage throughout the system affecting all levels and/or the connections between them. This second major assumption of the work, which has received a lot of criticism, is known as the globality assumption. This assumption contravenes the long held belief within neuro-psychology that the language processing system is modular and that damage may affect one level independently of neighbouring levels (e.g. Riddoch and Humphreys, 1994).

The resulting lesioned model was then parameterised to fit the data from a given aphasic individual by varying the degree to which damage to connection weights and/or decay rate is applied. The two lesion types lead to different error patterns, with connection weight lesions leading to more non-word and unrelated errors, and decay rate lesions leading to more related word errors. For a given individual a combination of a particular decrease to connection weight and a particular increase to decay rate should produce the pattern of overall severity of the naming deficit and pattern of error types. The severity of the damage required to each parameter will vary across the individual people with aphasia. The results of lesioning the model indicate that only certain error types vary with severity (formal, non-word and unrelated) while others do not (semantic and mixed) (see also data from Schwartz and Brecher, 2000, discussed in section 1.2.3.2).

The patient data consist of scores from the PNT from 21 people with fluent aphasia and include the number of correct items and numbers of semantic, mixed, unrelated, formal and non-word errors. People who made large numbers of no responses or descriptions were excluded from the study as the model is not able to cope with these data. Perseverative errors were not included in the model as each instance of naming is

treated in isolation, thus most instances of perseveration were coded as unrelated word errors.

For each patient's naming data a level of connection weight and decay rate was calculated which gave the closest approximation to the severity of the naming deficit and the error pattern. The model was then run 1,000 times for each patient and the best fit to the data selected. The degree to which the model matched the patient data was calculated by working out the root mean square deviation between the model's values and the patient's actual data. The authors claim a good fit for all patients. The root mean square deviation of each patient's data from the model varied however between 0.007 which represents a good fit, and 0.102, which represents a substantial discrepancy between model and patient. The best fit (patient LB) showed minimal discrepancy between model and patient, whereas the worst fit (patient WR) showed a large discrepancy between model and patient as shown in Table 2.2.

TABLE 2.2 Data from patients LB and WR (Dell et al, 1997)

	Correct	Semantic	Formal	Nonword	Mixed	Unrelated	RMSD
LB	.82	.04	.02	.09	.01	.01	
p = .007, q = .5	.82	.04	.03	.08	.01	.02	.007
WR	.08	.06	.15	.28	.05	.33	
p = .1, q = .94	.18	.09	.20	.37	.03	.13	.102

The columns show the scores for LB and WR and the model in terms of number correct and error types on the PNT. The first row for each patient shows their actual scores and the second row shows the model's best fit to these data. The final column gives the root mean square deviation (RMSD) between the two sets of scores. The values of p and q refer to the connection weight (p) and the decay rate (q), where normal p is 0.1 and normal q is 0.5.

2.5.3.1.1 Problems with this simulation

The model itself is set up on two sets of six words which is small by any account. It is difficult to extrapolate from this to the normal human's lexicon which typically consists of many thousands of words. The words are all CVC structure and it is possible that their neighbourhood properties are not generalisable to other words.

A second problem concerns the continuity hypothesis, and here consideration of the patient data will be instructive. The set of people who were included in the study performed well on the naming test: the range of scores was 94% to 8%, but the mean

score was 72% (standard deviation of 21). Thus the patients clustered around the top end of possible scores, with a few outliers performing poorly on the test. Only three people scored less than 50% on the PNT. Half of the patients scored 76% or more. Thus the model may well account for less impaired patients (those who are nearer to the normal population) but cannot tell us about more impaired performance.

This becomes clearer when the relationship between the overall success of an individual in the naming test is compared to the model's ability to fit their data. Analysis of the relationship between the individual's naming score and the RMSD of the model's fit to their data¹² showed a significant negative relationship between the two sets of values (Pearson's $R = -0.77$; $df = 19$; $p < 0.001$). This suggests that as naming ability increases so does the model's ability to simulate the data. The model can thus account for the patients who are nearer to the normal pattern of performance but does less well as success in naming increases in distance from the norm. The continuity assumption holds that aphasia is not qualitatively different from normal speech. The data here suggest that as the severity of the aphasia increases, the more the error pattern may diverge from that found in non-aphasic speakers.

A further difficulty with the simulation surrounds the number of cases excluded because of no response and circumlocutions. It could be claimed that participants were selected on the basis of the similarity between their profile and the pattern of normal speech errors. Furthermore the classification of aphasics' perseverative errors as unrelated word errors by default poses another question about the model's ability to handle all the aphasic data. This classification artificially negates the existence of such errors, whose very occurrence calls into question the continuity hypothesis.

Finally there are issues around the numerical treatment of the naming data. One example concerns the perseverative errors mentioned above: treating such errors as unrelated word errors artificially inflates the incidence of this latter error type. A further point to note concerns the data analysis in the study where raw scores for naming are analysed. As pointed out in Chapter One (section 1.2.3.2) different results may emerge depending upon whether one considers raw data or treats the data as proportions.

¹² This analysis was carried out by the author of this thesis and is not present in Dell et al's (1997) paper.

2.5.3.2 Foygel & Dell's (2000) semantic phonological model

One of the main criticisms levelled against the DSMSG model concerns the globality assumption built into the model. There is evidence from studies of people with aphasia to indicate that processing can break down in a discrete pattern. Yet lesions in the DSMSG model involve global damage to the system encompassing all levels of processing. In order to counter these criticisms and to render the model commensurate with patient data Foygel and Dell (2000) proposed a similar model to that described in 1997, but in which the connection weight parameter of the weight-decay model is divided into a lexical-semantic and a lexical phonological component. These two weights could now be independently lesioned, thus allowing the model to resemble more closely the patterns of deficit seen in individuals with aphasia. A given individual will have more or less of a semantic-lexical connection weight lesion and more or less of a lexical-phonological connection weight lesion. Decay rate is not lesioned in the new simulation. In all other respects the two models are the same.

The model was parameterised using the same method as that used for the DSMSG model. The best fit for each of the 21 patients described in Dell et al (1997) was obtained and these were compared to the best fit provided by the original model. Although the two models incorporate different types of damage it is instructive to note that both produce similar outputs thus capturing the data well. This is important as the first model's globality assumption is deemed a weakness. The second point to note is that the two models reveal counter-intuitive effects of damage. One striking example is shown in the semantic-phonological model's production of *formal* word errors as a result of semantic-lexical damage. This is explained in terms of damaged semantic-lexical connections, which are not sufficiently strong to counteract the effect of the activation from phoneme-lexemes, thus formal (unrelated semantically) words are more likely to be selected over semantically related neighbours (which are receiving weakened activation).

Two sets of statistics are provided: the root mean square deviation (as used in DSMSG) and chi square. Using RMSD the semantic-phonological model provides a better fit for 11 of the 21 patients; using chi square the semantic-phonological model provides a better fit for 15 of the 21 patients. The authors go on to demonstrate that although both models provide similar accounts of the data to a similar degree of accuracy, when the

error space of the two models is considered (in six-dimensional space) and the patients are located within this space, the semantic-phonological model provides a better account of the patients with no patients falling far outside the space (whereas for the DSMSG model patient WR's profile falls markedly distant from the error space).

All of the criticisms levelled at the DSMSG model apply to this new model, apart from the globality assumption. What these two studies demonstrate is that computer simulations can produce some of the patterns of aphasic naming albeit within certain tight constraints, and that new hypotheses regarding the source of errors may emerge from such studies, for example the occurrence of formal errors following semantic-lexical damage. This may lead aphasiologists to reconsider testing procedures and the interpretation of test results. At present the models fail to account adequately for all the data and consider only a small selected group of people with relatively mild naming deficits, and further simulations are required to develop these ideas further.

2.6 SUMMARY

In this chapter the history and background to current methods used in the study of language after brain damage has been outlined. Some of the key assumptions of cognitive neuropsychology have been discussed. The various sources of data which have been called upon to develop theory in spoken word production including speech errors, tip of the tongue data, reaction time experiments and data from people with aphasia have been discussed. Important theories in this area have been described and the strengths and weaknesses of each outlined. The degree to which each can account for the patterns of spoken word production seen in aphasic word finding has been analysed.

This necessarily lengthy account has been provided in order to depict the background against which the investigation of the spoken language production of individuals with aphasia currently takes place. Theories of spoken word production must be able to account for the patterns of retained and impaired processing found in any given individual, and ideally, for the patterns of recovery witnessed in such people. The data presented in this thesis will be considered in the light of the findings outlined in this chapter (see Chapters Seven, Eight and Nine).

CHAPTER THREE: SPOKEN WORD PRODUCTION IN CONVERSATION

3.0 INTRODUCTION

In the previous two chapters a review of what is known about word retrieval in aphasia and theories of spoken word production and their ability to account for the data from aphasia was presented. Most of the data considered in those chapters consists of single word production in experimental conditions. Most naturally occurring speech happens in conversation however. One of the aims of this thesis is to explore whether the effects of therapy can be measured in terms of changes to conversational behaviour. For example, does someone produce more nouns in conversation after therapy, and if such changes occur, do they affect the conversational interaction. In order to investigate this issue responsibly recourse to current thinking in the area of Conversation Analysis (CA) is warranted. It was deemed appropriate to base the analysis of the interactional effects of therapy on constructs gleaned from CA. In this chapter an outline of CA is provided along with a discussion of what this form of analysis has contributed to our understanding of aphasia. This is then followed up in Chapter Six with a description of a means of measuring therapy effects on conversation.

3.1 BACKGROUND

Although the past hundred years have seen major advances in the understanding of linguistics and psycholinguistics, it remains true that most research carried out in this area considers speech in non-interactive tasks. These are often conducted in a laboratory setting (see details of such research in Chapter Two), or in the case of linguistics in naturally occurring speech but as isolated phenomena. Relatively little theory exists concerning language production in conversation. One of the reasons for this lack of research concerns the number of variables involved in measuring conversation: it is relatively easy to measure word production in a laboratory setting, whereas it is extremely difficult to measure the same phenomenon in naturally occurring conversation. For example, it is not possible to control key variables in conversation, such as word frequency. Thus most of what is known about language production concerns the act performed in laboratory conditions involving non-interactive constrained speech tasks such as single word production.

Conversation Analysis (CA) provides new insights into conversation, by viewing it as a collaborative process, through which mutual understanding (inter-subjectivity) is achieved. Through taking turns in a conversation the speaker and their partner co-construct meaning. The two parties involved are jointly responsible for the outcome of the interaction. It is worth noting that the early proponents of this form of inquiry were sociologists and not linguists. Much of what is written about conversation consequently disregards the language used and concentrates instead on the sequential acts performed by the two speakers. CA can offer a way of investigating the effects of aphasia on interaction, but offers no insights into the linguistic demands or impairments impacting on the conversation. CA describes the result of the aphasia on the conversation e.g. a word finding difficulty may cause breakdown and subsequent collaborative repair. It will not interpret any further beyond this¹³.

This approach is attractive to aphasiologists as it considers the speaker in the real life setting, and thus has great ecological validity. Moreover the notion of collaboration removes the glare of the spotlight from the aphasia, and considers instead the two speakers and the effect of the turns they produce on the flow of the conversation. Many studies in recent years have focussed on the specific nature of the conversations between aphasic speakers and their primary speaking partners (e.g. Milroy and Perkins, 1992; Lesser and Milroy, 1993; Ferguson, 1992; Ferguson, 1994; Perkins, 1995; Wilkinson, 1995). From this an intervention may emerge which targets the non-aphasic partner's actions in conversation, rather than the aphasic person's disability (e.g. Booth & Perkins, 1999).

3.2 CONVERSATION ANALYSIS

Conversation analysis is a method of analysing naturally occurring conversation, and does not form a theory in itself. CA originated in the work of the sociologists Harvey Sacks and Emanuel Schegloff. They approached conversation from an ethno-methodological background, which emphasises data over theory, and attempts to describe and derive truths from the observed data. This approach has no a-priori assumptions. Naturally occurring conversations are transcribed and studied in order to identify patterns within the interaction. CA does not generally involve quantification,

¹³ Indeed for many proponents of CA the assumption that a word finding difficulty is occurring is an assumption too far, unless this is made explicit by the speaker involved.

although in drawing out generalities some recourse to quantification is implicit. The overriding concern of CA is to describe the nature of turn taking in conversation and the organisational principles that underlie the apparent disorder witnessed in the interwoven speech of several speakers.

3.2.1 Sequentiality

A key notion within CA is that of sequentiality. This refers to the general finding that conversation between two speakers involves a series of turns, which are not merely temporally ordered, but are organised into sequences. Speakers share an orientation to the possible ways in which sequences are constructed. The most obvious example of this is a question and answer sequence. For example, in the following sequence involving PH who is one of the people with aphasia taking part in this research, and her chosen conversation partner SP, SP's question about the chiropodist in line 1 is responded to fully by PH in lines 2 and 4. See Appendix One for a synopsis of notations used in the transcriptions.

Conversation Sample 3.1

- 1 SP when's he seeing you then
- 2 PH ah, I'm gonna phone up=
- 3 SP =have you
- 4 PH every six six weeks I've got to go. got to phone them up
- 5 SP hmm

PH = person with aphasia

SP = conversation partner

These pairs of turns, in which production of the first part of the pair requires the production of the second part (or one of the range of possible second parts) by the next speaker in serial order, are known as adjacency pairs. Other examples include greetings and responses, invitations or offers and their acceptance or rejection. The pairs need not be serially adjacent, but the production of the first part by the first speaker requires, at some point, negotiated by the two speakers, the production of the second part by a second speaker.

The claim that speakers orient to adjacency pairs is supported by the observation of the first speaker when the next speaker has failed to produce the second part. In the next example (Hutchby and Wooffitt, 1998: 42) the parent fails to produce the required second part (an answer to the question).

Conversation Sample 3.2

- 1 C Have to cut these Mummy. (1.3) Won't we
- 2 Mummy
- 3 (1.5)
- 4 C Won't we
- 5 M Yes

C=child

M=mother

From Hutchby and Wooffitt (1998:42)

In this example the child asks her mother a question in line 1, to which her mother fails to respond in the 1.5 second pause in line 3. The child calls attention to this violation by repeating the question in line 4.

3.2.2 Turn taking

Given that conversation must involve at least two people, it follows that the speech produced is shared between those speakers. This sharing of speech contributions is known as turn taking. Within conversation turns can vary freely along three parameters: the turn form, the turn content and the turn length (Hutchby and Wooffitt, 1998). In everyday conversation speakers are able to produce turns rapidly with little gap or overlap. The issue of how this is achieved is now considered.

3.2.2.1 Turn construction units and transition relevance points

What constitutes a turn remains a matter of some debate. The term *turn construction unit* (TCU) is used to denote a contribution to a conversation by one speaker to which the speakers orient in planning subsequent turn-taking in the conversation. When a TCU is concluded another speaker or the same speaker may begin another turn construction unit (depending upon the nature of the first TCU). A turn construction unit may be a

non-verbal sound or an action, a word, a fragment, a sentence or a series of sentences. A change of speaker happens in conversation at places where speakers project a transition relevance point (TRP). This is a point at which a change of speaker may occur and evidence for speakers' shared awareness of a possible TRP is shown by the manner in which speakers predict the end of a TCU and the advent of a TRP. The prediction (or projection) amounts to an anticipation of the end of the turn, based on linguistic and supra-segmental information. In English this might involve anticipating the end of a syntactic structure, a fall in intonation in a statement, dropping of eye gaze (e.g. Goodwin, 1981), or a movement of the hands or the whole body. The first speaker may signal a TRP clearly, as in sample 3.1, line 1 and sample 3.2, line 1, or the second speaker may anticipate it.

Where a second speaker is anticipating the completion of a turn and misprojects this, overlap may occur. This is particularly the case where tags are used at the end of a statement or question. In the latter instance a degree of overlap may occur, as is shown in PH's conversation with SP (sample 3.3).

In this example PH predicts a transition relevance point at the end of SP's utterance "you were lucky then" (line 1). SP adds a tag to the main sentence however ("weren't you") with which PH's turn (agreement with the proposition offered by SP in line 1) overlaps. Overlap is tolerated in this instance, as is seen by SP's turn in line 3 which offers a comment on PH's agreement in line 2, even though this agreement fell in the period of overlap and could have been ignored by SP.

Conversation Sample 3.3

- 1 SP you were lucky then [weren't you
- 2 PH [yeah
- 3 SP oh good
- 4 PH yeah
- 5 SP yeah
- 6 SP and did you stay at the hotel

PH = person with aphasia

SP = conversation partner

Sacks, Schegloff and Jefferson (1974) studied conversational turn taking extensively and derived from their observations a set of rules of turn taking. One example of these is that if a first speaker completes a turn and no second speaker is indicated, any of the speakers may take the next turn. In some studies of CA taking a turn when not explicitly required to do so by the rules of turn taking has been labeled 'initiation' (see 3.2.2.2 below). In various studies since then validation of the claims made by these rules has been found. The rules outline how turns are negotiated between speakers.

3.2.2.2 Types of turns

In peer conversations involving speakers in a social interaction (as opposed to an institutional interaction), speakers' roles are negotiated rather than specified¹⁴.

Depending upon the topic and the sensitivity of the speakers to each other's needs, one or other speaker may dominate the conversation. Alternatively equal participation may emerge. This can be analysed by looking at the types of turns produced by the speakers. The terms initiation, response, and minimal response have been used to differentiate the interactional consequences of different types of turns. Unfortunately there is little agreement about what these constitute, and few studies have looked at the reliability of the terms. Briefly however, an initiation may occur at a TRP where no next speaker is selected (Sacks et al, 1974). A response occurs where a next speaker has been selected by the first part of an adjacency pair (question/request/proposal must be followed by response). A definition of minimal responses is hard to find. For some researchers this includes all turns which serve to hand the floor back to another speaker.

In sample 3.4 SC (a person with aphasia) is talking with his wife LC. In lines 2 and 3 there is an example of a question and response. In line 3 SC initiates a turn. He has completed a turn in line 2 (the TRP is marked by the completion of a syntactic unit). No speaker is selected for the next turn, so SC's turn in line 3 amounts to an initiation. Lines 5, 6 and 7 consist of minimal responses. In CA the term 'discourse markers' is

¹⁴ In institutional settings (where one person is a professional, the other a lay person) the speakers' perception of their relative power and their usually mutual adherence to a preordained (culturally specified) agenda often determines the turns the two people will produce.

Conversation Sample 3.4

- 1 LC anything else you want to do when you've done that
- 2 SC emm. there is nothing in my mind (2.61 seconds)
- 3 em. there is a little exercising
- 4 LC you're gonna do some exer [cising
- 5 SC [yes
- 6 LC right
- 7 SC yeah

SC = person with aphasia

LC = conversation partner

used to refer to “minimal speech items whose chief function is to organize and structure interaction” (Lesser and Milroy, 1993: 218). One example of this is the use of ‘mm’ to signal continued attention to another’s speech. Another function of minimal responses is to hand the floor back to the other speaker. For example in sample 3.5, KR (who has aphasia) initiates a second turn (line 3) after her husband SR’s minimal response (line 4), and her husband signals his continued attention with “mhm” (line 4) as KR searches for a word in line 3.

Conversation Sample 3.5

- 1 KR the children are ehm. the children are alright=
- 2 SR =ok
- 3 KR ehm. Sam is [ehm. eh. school
- 4 SR [mhm

KR = person with aphasia

SR = conversation partner

3.2.3 Repair

Things can go wrong in conversation. For example speakers make errors in sound and word selection, and they make errors in turn projection leading to unmanageable overlap. Hutchby and Woollfitt (1998: 57) refer to the phenomena labelled repair as

“repair of the turn-taking system”. Boles and Bombard, (1998: 550) describe repair as “an attempt to modify one's own or the other person's utterance for the purpose of clarification”. The following example shows an error in word production from PH's conversation with SP.

Conversation Sample 3.6

- 1 PH something with /tu/ . T . /teləsu/ or something .
- 2 I don't know
- 3 SP Saint Lucia
- 4 PH yeah
- 5 SP Saint Lucia
- 6 PH yeah something like that [mm
- 7 SP [ooh that was nice then

PH = person with aphasia

SP = conversation partner

In this example the usual turn by turn sequence is inhibited PH's inability (line 1) to complete her turn due to a difficulty in producing the correct word form. The conversation is then diverted into sorting this out, and this constitutes the repair.

Discussion of the topic does not resume until both speakers accept the resolution of the problem. In this extract PH is struggling to produce the name of a place and makes a partial attempt and a phonologically related attempt at her target (line 1). SP offers a possible target (line 3) which PH appears at first to accept (line 4) and the repair appears to be completed. SP then offers the target word again (line 5) as the culmination of the repair. PH's subsequent turn (line 6) indicates that this is not the resolution and that the selected target is not right. SP's subsequent turn (line 7) rounds off the sequence by ignoring PH's demurrals in line 6 and returning to the topic (PH's grandson's honeymoon in an exotic place). This example demonstrates how people with aphasia may accept being misunderstood within conversation, lacking the resources to repair the breakdown effectively, or judging that further repair is the dispreferred option in the interests of continuing the conversation.

Although breakdown in normal conversation does occur it is generally resolved within three turns (Schegloff, Jefferson and Sacks, 1977). Moreover, there is a preference

within conversation for speakers to repair breakdowns themselves within the current turn (so called self-initiated self-repair).

Conversation Sample 3.7

- 1 SP was she the only child
- 2 PH no his his . girl /kɛɪ/ er boy came in from abroad
- 3 SP oh did he

PH = person with aphasia

SP = conversation partner

This is preferred for a number of reasons one of which has to do with saving face amidst the socially delicate business of handling incompetence in speaking. This occurs within aphasia as well. In sample 3.7, PH carries out self-initiated-self-repair. She makes a semantic error (labelling boy “girl”) in line 2 which she rapidly corrects.

3.3 CONVERSATION ANALYSIS AND APHASIA

3.3.1 Turn taking in aphasia

People with aphasia have difficulties in processing speech and in producing speech both in constrained tasks and in conversation. In the latter situation language processing deficits can lead to disruption to the flow of conversation. The person with aphasia may fail to understand or be slow to understand a previous turn and thus be unsure how to continue; they may fail to take up a turn at a transition relevance point because of difficulties initiating production or including difficulties finding the right words; they may also fail to complete a turn due to word finding difficulties, or syntactic processing problems. People with aphasia do take turns however and do not on the whole violate the rules of turn taking. The quality of turn taking may differ from that of non-aphasic speakers, or from that of the pre-morbid self. The fact that turn taking is occurring may obscure the possibility that the turns produced by the person with aphasia differ from those of the conversation partner in terms of the burden of the conversation taken up, the amount of contribution to the conversation, and the ability to introduce a new topic, or maintain a topic.

Conversation Sample 3.8

- 1 DS Eirean's in instead today
- 2 DA oh yeah yeah yeah tis yeah
- 3 DS cos she's gone to a funeral
- 4 DA oh yeah
- 5 DS I dunno whose
- 6 DA mm yeah

DA = person with aphasia

DS = conversation partner

Examples of how turn taking may be disrupted in aphasia are now described¹⁵. These are generally due to linguistic breakdown, but the consequences of the breakdown will differ. One common consequence involves the person with aphasia relying on minimal contributions to the conversation.

Sample 3.8 is from DA and his wife DS. DA has a severe expressive aphasia. In this extract DS gives a description of who is working with her that day, and the events precluding another worker's attendance (line 3). DA's contributions demonstrate his adherence to the turn taking rules, but his turns can serve only to signal his attention and interest, and he is unable to add semantically to the topic. Lesser and Milroy (1993) term this sort of minimal turn production the 'perverse passive', a turn which serves only to hand the floor back to the other speaker. In many cases, this form of turn is the only option available to the person with aphasia, and may be used strategically to enable them to take part in the interaction at a level that is manageable for them. Other instances where a person with aphasia does not participate fully include an inability to initiate turns, an inability to signal the intention to initiate a turn, not responding when selected, long pauses which are intolerable to the other speaker, who will then retake the floor, or a difficulty holding the floor when taking a turn. An example of an increased tolerance to pauses in the conversation partner is shown in sample 3.9, where PH and SP are discussing where PH's daughter has gone. Here however SP waits for PH to find the word she is looking for (target = 'Butlin's') after a 2.5 second pause in line 2.

¹⁵ It is worth noting that the behaviours outlined in this section are also found in non-aphasic speech.

Conversation Sample 3.9

- 1 SP where's Carol today
- 2 PH she's down at (2.5 seconds) Butlin's
- 3 SP oh yes

PH = person with aphasia

SP = PH's friend

Conversely a speaker with fluent aphasia may fail to relinquish the floor. By failing to pause, or to signal the end of the turn through eye gaze or body movement, the person with fluent aphasia may fail to signal any transition relevance points to the other speaker. Thus the other speaker's attempts to take the floor are not registered by the first speaker.

3.3.2 Aphasia and repair

The usual pattern of repair in normal conversation is self-initiated and self-completed (self-initiated self-repair, SISR). Aphasic individuals do carry out self-initiated-self-repair, which may involve partial attempts at a word, rephrasing, pauses and fillers (see conversation sample 3.7, line 1 above). Conversations between a non-aphasic speaker and an aphasic speaker differ from those between two non-aphasic speakers because, when a breakdown arises, the aphasic partner may well be unable to solve the problem alone, for example in searching for a particular word. In such instances considerable participation from the non-aphasic partner occurs. The result is sequences of talk in which both parties seek a successful resolution to the difficulty, by a variety of means, often extending over long stretches of conversation.

Oelschlaeger and Damico (1998) refer to repair in aphasia as 'joint collaborative sequences'. They prefer to view these sequences as a means of interacting which form an act of communication which is valuable, regardless of the normal criteria for success. They refer to joint productions, which involve "the initiation of a turn by one speaker and the syntactically and semantically coherent extension or completion of that initiation by another speaker." (Oelshlager and Damico, 1998: 461). Such examples also occur in non-aphasic speech, for example Sacks (1992) refers to joint productions and

utterance completions in non-aphasic speakers conversations. Laakso and Klippi (1999) report that in conversations where there are aphasic participants, problems are often treated sequentially in a collaborative fashion.

Clark and Schaefer (1987) describe a principle of 'least collaborative effort' inherent in normal conversational repair. A conversation participant chooses the strongest initiator of repair consistent with her or his current state of understanding. For example in the following exchange: A: Every time he comes that happens. B: What does? B uses all the information she has to enable her to question directly the part she lacks information about. This principle pertains to conversations between non-aphasic speakers but is often violated in conversations involving people with aphasia. In such conversations there are instances where the non-aphasic partner persists in questioning the person with aphasia, when he or she knows what they are trying to say but is dissatisfied with the turn they produced. This may be because the aphasic speaker failed to produce a target word, or produced an error. The subsequent repair, initiated by the non-aphasic speaker, involves an attempt to get the person with aphasia to achieve the correct production, and may involve strategies such as phonemic cueing by the non-aphasic speaker. This may constitute an attempt to allow the person with aphasia to practise talking. Sample 3.10 demonstrates this point. PH and SP are looking at PH's old family photographs.

Conversation Sample 3.10

- 1 SP what's Ben to me
- 2 PH your Ben
- 3 SP yeah . what is he to me
- 4 PH now you've got me . erm .erm
- 5 SP what's your mum and dad .
- 6 what's your dad to your mum (2.1 seconds)
- 7 SP her
- 8 PH my mum and dad
- 9 SP yeah it was your mum's what
- 10 your dad was what to your mum
- 11 PH my mum
- 12 SP yeah . her
- 13 PH /hʌz/ . [erm . father

- 14 SP |yeah
 15 no
 16 PH /hʌzbəlænd/
 17 SP yes husband you've got it
 18 PH husband

PH = person with aphasia

SP = conversation partner PH's friend

In this extract SP leads the conversation into lengthy collaborative repair such as would never be seen in a peer conversation between non-aphasic people. PH signals clearly that she cannot find the word for 'husband' (line 4). This failure is then highlighted by SP through her repeated endeavours to cue PH into production of this target, which is ultimately successful (lines 16 – 18). Although such sequences do not form part of normal interaction, and indeed would be embarrassing were they to, they do form part of many dysphasic people's interactions with non-aphasic speakers. It is clear also in many cases (although not all) that little irritation or annoyance is displayed by the person with aphasia, signalling their compliance with this form of interaction.

Repair in aphasia differs also from that in non-aphasic conversation in the way in which it is concluded. Milroy and Perkins (1992) compare the endings of repair sequences in aphasic conversations to the endings of normal conversations, where speakers make various moves in sequential fashion to close down the sequence. This is an unusual feature of aphasic repair which is not found in non-aphasic conversations, where repair is usually dealt with rapidly without calling attention to it.

3.3.2.1 Causes of breakdown in aphasic conversation

For many people with aphasia word finding difficulties form a significant part of their language disability. Lesser and Algar (1995) and Perkins, Crisp and Walshaw (1999) found that lexical retrieval problems were the most common trouble source for most of the people with aphasia they studied. In addition to lexical retrieval problems they found a number of overt speech errors including jargon, neologisms, phonological paraphasias, semantic paraphasias, filled pauses and agrammatic breakdown in production.

3.3.2.2 *The nature of repair in aphasic conversation*

Ferguson (1996) reports that people with aphasia show 'trouble-indicating behaviour' (Schlenk et al, 1996) which is followed by hypothesis testing where the partner offers candidate targets as guesses. Lubinski, Duchan and Weitzner-Lin (1980) describe 'hint and guess' sequences occurring between the person with aphasia and their spouse, or between the person with aphasia and the speech and language therapist. An example of this is shown in sample 3.11. SC is trying to retrieve the name of an English stately home (Blenheim) and, unable to do so, offers LC information about the place. In the many turns that follow (the sample here is a small extract from a repair sequence stretching over five minutes of conversation) LC attempts to guess at the target.

Conversation Sample 3.11

- 1 SC if we look at the most famous of all we'd never get in . it cost the earth
- 2 LC why is that (3.3)
- 3 SC oh . eh dear . of a summer it's thick very expensive we're . oh alright I'll try
and explain to you . em they are uh they are known it's (3.4)
- 4 uh most expensive
- 5 LC yes
- 6 SC often you get there . oh (sighs) this is difficult isn't it (3.60)
- 7 LC hmm
- 8 SC [hehehe
- 9 LC [well you said in the country=
- 10 SC =right
- 11 yes it's a /bʌn/ and it's a place
- 12 LC yeah
- 13 SC a large area . a place we go . everywhere is expensive
- 14 LC lots of places are [expensive
- 15 SC [no this is special=
- 16 LC =eh I can't think where you mean

SC = person with aphasia

LC = conversation partner

Laakso and Klippi (1999) describe people with aphasia attempting self repair initially. If this fails the person with aphasia establishes a collaborative framework by clearly shifting the orientation to a co-participant. They identified a number of phases in repair in conversations of people with aphasia: i) problem establishment; ii) establishing a collaborative co-participation framework; iii) hint and guess phase; iv) long confirmation phase. People with aphasia, like non-aphasic people, tend to prefer self-initiated self-repair and will attempt this first however.

Wilkinson (1995) reports that breakdown can occur long before repair is initiated - a repair will begin only when it is clear that there has been a misunderstanding, the cause of which can be traced back several turns.

3.3.2.3 Influence of partner

Ferguson (1994) found more instances of repair in conversations between non-aphasic and aphasic speakers than in conversations between two non-aphasic speakers, and more 'other' repair in conversations involving aphasic speakers. Perkins (1995) described differences in collaborative repair in the conversations of the aphasic participant EN. Her conversation partner rarely engaged in collaborative repair whereas the researcher did (described further in Milroy and Perkins 1992). Clearly there are differences across lay and professional conversation partners. The two differ in the amount of knowledge they share with the person with aphasia, the friend or relative usually having more knowledge of personal issues, and of the person with aphasia's speaking style. Perceptions of power and authority will affect both speakers in a lay-professional interaction. These factors need to be taken into account when considering conversation samples.

3.3.2.4 Types of repair in aphasia

Lesser and Algar (1995) described the ways in which repair is initiated in their study of the conversations of two people with aphasia each with a non-aphasic speaking partner. These are as follows:

- the person with aphasia asked for help and the pair jointly reached the target;
- the conversation partner asked for and received clarification;
- the conversation partner, knowing the target, corrected a mispronounced utterance;

- the conversation partner's guess is not resolved;
- the conversation partner's request for clarification is not resolved.

In addition to these occurrences Wilkinson (1995) and Ferguson (1992) describe the conversation partner being asked to find a word for the person with aphasia, and Lubinski et al (1980) describe the person with aphasia hinting at the target to support the conversation partner in their guesswork.

3.4 RELIABILITY AND THE MEASUREMENT OF CONVERSATION

One overwhelming difficulty in this line of investigation concerns the unconstrained nature of conversation, with two conversations between any two speakers showing variation across a number of parameters. Boles and Bombard (1998) report an extreme example of this variation in their study of repair in conversation. In a set of 11 conversation samples between one aphasic person and their speaking partner, occurrence of partner-initiated repair ranged from seven to 64 instances. Such variation demonstrates that quantification of this type of data is potentially unstable, thus conclusions about change cannot easily be drawn. As a result the use of conversational data as an outcome measure is problematical. Both qualitative and quantitative differences found between any two conversations may be simply part of the inherent variation. Additionally many aspects of conversational data are open to interpretation by the analyst (see comments in section 3.2.2.2 and further discussion in Chapter Six).

A number of studies have attempted to establish the reliability of conversational data. In most cases a percentage of the data was analysed by two researchers (e.g. Watson, Chenery and Carter, 1999; Oelschlaeger, 1999; Boles, 1998; Boles and Bombard, 1998). Such studies have looked at agreement of the accuracy of the transcription (e.g. Ferguson, 1996); incidence of trouble-indicating behaviours, repair trajectory and repair type (Watson et al, 1999), and repair sequences (Oelschlaeger, 1999). In most cases inter-rater and intra-rater agreement was sought. These are one form of reliability considering the same piece of data. The potential variation across conversations necessitates a measure of test-retest stability in addition to inter and intra-rater reliability.

3.5 SUMMARY

In this section the impact of aphasic language deficits upon conversational turn-taking, and the means by which repair is conducted, have been reviewed. In order to make a real impact upon a person's life therapy should ideally affect everyday interaction. A word finding difficulty can have a serious impact upon conversation, and one way of addressing this in assessment and therapy is to analyse the consequences of the difficulty on specific aspects of conversation interaction. In order to do so the reliability and stability of the aspects in question needs to be established. These issues will be addressed in Chapter Six in considering a possible quantitative measure of conversation.

CHAPTER FOUR: RECOVERY AND REHABILITATION

4.0 INTRODUCTION

In this chapter recovery and rehabilitation after a stroke are discussed. A principled approach to rehabilitation is outlined which provides a coherent account of the relationship between specific stimulation such as is used in language therapy, and the effects of that on brain plasticity. This is followed by an account of how theories of language processing have informed facilitation and therapy interventions. A review of phonological therapy is provided as this forms the first phase of therapy used in this study. Finally approaches which combine specific stimulation of word sets with communicative interaction are described.

4.1 MECHANISMS OF RECOVERY

4.1.1 Cortical plasticity

4.1.1.1 Neuronal regeneration and cortical plasticity

It was a strongly held belief throughout the early twentieth century that a brain once damaged could not regenerate neuronal tissue. A second strongly held belief was that an adult human brain shows little or no capacity for reorganisation of function through cortical plasticity. Robertson and Murre (1999) claim however that there is now significant evidence for what they term “experience-dependent synaptic changes” (Robertson and Murre, 1999: 545) in the non-damaged human brain, involving modification of synaptic connectivity, dendritic arborisation, and axonal sprouting. Significantly they base their theory of recovery on the premise that principled rehabilitation should maximise the residual capacity in the damaged brain for plastic reorganisation of brain systems. This has two implications: first, after insult neuronal tissue may re-grow and severed connections be restored; and second, the operation of certain mental functions may transfer from the damaged area to a related undamaged area¹⁶. Both processes are strongly implicated in the recovery or restitution of function.

¹⁶ This claim has been current in the neuropsychological literature for some time e.g. Coltheart (2000) has argued that the (pathological) reading pattern seen in deep dyslexia results from the undamaged and unspecialised right hemisphere’s involvement in reading after brain damage to the left hemisphere.

Evidence of cortical plasticity is seen in normal adult learning where changes within cortical sensory and motor fields are witnessed in the context of specific learning experiences (see e.g. Keefe, 1995; Kolb, 1996). Conversely, lack of appropriate stimulation of the non brain-damaged individual can lead to loss of connectivity: “there is abundant evidence that even in the undamaged brain loss of stimulation or disuse results in declining connectivity within a circuit” (Robertson and Murre, 1999: 551). This statement has important implications for intervention after brain insult.

Hebb (1949) proposed that a strengthening of synaptic connections occurs when pre and post-synaptic neurons are co-active. Two neurons or two groups of neurons which have been disconnected by damage may become reconnected if activated at the same time, giving rise to the maxim ‘cells that fire together wire together’. This principle known as ‘Hebbian learning’ underlies the major theories of cortical recovery and has implications for the implementation of language and other cognitive therapies. A warning note is warranted at this point however. Many of the claims regarding cortical plasticity come from animal research, and those investigating human recovery focus on the whole on motor and visual perceptual functions. It is unclear at this point whether these findings will generalise to the higher cognitive domain of language.

4.1.1.2 Degree of damage and implications for rehabilitation

In a review of the processes of rehabilitation and recovery Robertson and Murre (1999) distinguish three levels of damage. Where there is *mild* damage they propose that autonomous recovery will occur without intervention. Where there is *severe* damage with substantial loss of connectivity there will be a failure of recovery. One of the challenges for rehabilitation professionals is to identify this situation and act appropriately, by not trying to restore functions which are not susceptible to rehabilitation efforts. The rehabilitation professions are at present far from being able to judge this with any degree of accuracy.

The third level of damage is *moderate*: “Circuits that have lost an intermediate number of connections may be in a critical state where they could either lose connectivity completely or alternatively recover patterns of connections sub-serving the impaired neuropsychological function. Rescue versus collapse may depend on such circuits receiving precisely targeted stimulation fostering Hebbian-based reconnection of the

partially disconnected network” (Robertson and Murre, 1999: 551). It is exactly this “precisely targeted stimulation” that therapy for word finding deficits in aphasia should be trying to achieve.

Of these three patterns, people with a mild lesion will not require rehabilitation, those with a severe lesion will require compensatory methods, and those with a moderate lesion will benefit from directed stimulation. In this latter group there is thought to be sufficient neural connectivity and mass for cortical plasticity to play a part. Cortical plasticity is maximised by directed experience. To this end rehabilitation needs to be scientifically based, that is to identify appropriate targeted stimulation for a given deficit in function. Within the realm of language disorder, this means using current models of language processing to assess and plan intervention. Moreover, the timing and intensity of the intervention is of paramount importance, if disuse can result in a decline in connectivity and hence in function.

4.1.2 Connectionist accounts of recovery

Connectionist accounts of language processing offer an opportunity to study the possible mechanisms involved, the operation of variables such as word frequency on language processing, and the recovery of the system after damage has been implemented. A number of studies have reported ‘lesioned’ computer simulations of a particular language function, and studied the resulting behaviour of the model in terms of patterns of breakdown and of recovery. Much of this work has been carried out in the field of acquired dyslexia (e.g. Plaut, McClelland, Seidenberg and Patterson, 1996).

Dell et al (1997) looked at the capacity of their model of single word production to regain function after damage. They compared the model’s performance with different types and amounts of damage to the patterns shown by a group of people with aphasia. The DSMSG model was set up and then lesioned in terms of decay rate and connection weight, to get the closest fit between the model and the performance of 21 people with aphasia on a test of picture naming. These people were later (between one and a half and nine months) retested on the same picture naming test, thus providing a measure of recovery in terms of numbers correct and pattern of errors. In the model the process of recovery was simulated by the gradual return of the connection weight and decay rate parameters to their normal values. If the model provides a valid account of impaired

naming, it should be able to simulate the recovered as well as the initial patterns of impairment of the individual patients. Moreover, the character of the fit for any given individual should not change, for example a person whose naming was best modelled by a connection weight lesion should not recover into a decay rate lesion.

The recovery part of the study involved ten of the 21 people with aphasia. The authors claim a good fit for all ten people from the model. The character of the fit was maintained for eight of the people. That is, someone with a primarily decay lesion maintained this pattern in the course of recovery. Of the remaining two, one recovered function to a normal level, and the final person moved from a high weight lesion to a low weight lesion. As his recovered performance was at near normal levels the authors do not consider this a problem.

Studies into the recovery of spoken word production using connectionist models are at this point rare. Dell et al's (1997) and Foygel and Dell's (2000) investigations into this area represent the start of a major research enterprise. There are however significant limitations to the research conducted so far. One limitation concerns the small number of vocabulary items the models are trained on. A second concern relates to Dell et al's (1997) model which implemented global damage to the system. This is at odds with findings of discrete damage to one processing level within the cognitive neuropsychology literature.

4.2 PSYCHOLINGUISTIC MODELS AND PRINCIPLED THERAPY

Attempts at guided recovery in the field of word finding deficits in aphasia have developed in the past 20 years as a result of the emergence of robust models of normal spoken word production, against which to formulate a theory of how breakdown occurs in aphasia and hence how intervention may affect its recovery. The studies published in this field have investigated the effects of different types of stimulation upon word finding. Howard (1985a) distinguished between cues, facilitation and therapy. A cue is the administration of one stimulus upon the word-finding breakdown and the measurement of its immediate effects; facilitation refers to the maintained effect of the cue over time; therapy is the term used to refer to repeated administration of cues over time. The studies reviewed here are selected according to their use of "guided recovery" techniques. That is, a set of target words are treated with a principled intervention

technique, in line with theories of spoken word production, and the effects of the stimuli on spoken word production are measured over time.

4.2.1 Effect of cues on spoken word production

The motivation behind the methods used in facilitation and in therapy comes from the single step model. As this proposes only semantic and phonological levels, these are the two forms of intervention trialled so far. Within the rubric of phonological cues fall a number of methods: provision of the first phoneme, provision of the whole word (repetition), and provision of the word rime (in CVC words this would include the vowel and the final consonant). A more recently implemented form of cueing involves the use of progressive cues, where the person is exposed to more and more of the target word's phonology, until successful word production is achieved (e.g. Lambon Ralph, Cipolotti & Patterson, 1999; Hickin et al, 2002; Nettleton and Lesser, 1991). This form of cueing was used in the therapy described in this thesis.

4.2.1.1 Evidence from normal participants (priming studies)

There is significant evidence from a number of studies that the implementation of specific cues can facilitate spoken word production in non-aphasic participants. Semantic priming and repetition priming are the two main techniques used, and most of the research has focused on word recognition rather than production. Seidenberg et al (1984) found speeded word recognition when a target is immediately preceded by a semantically related prime. This facilitation is not mirrored in production tasks however. Wheeldon and Monsell (1994) found an inhibitory effect of an immediately preceding semantically related prime on picture naming. They proposed that the prime word is in competition with a co-activated target word.

Repetition priming has been shown to have robust and long lasting effects on word recognition and on picture naming. Cave (1997) found a lasting effect of repetition priming on word production up to 48 weeks after administration of the prime.

Repetition of a target word can also facilitate word production in aphasic speakers. Understanding how this occurs may help to identify people for whom it will be an effective form of intervention. Monsell and colleagues investigated repetition priming effects and their locus in relation to models of speech production. Wheeldon and

Monsell (1992) looked at single word production in response to hearing a definition, and investigated reaction times on subsequent naming. They found reduced reaction times for naming at lags of up to 90 items. Two manipulations allowed them to hypothesise the locus of the priming effect. They argue that the effect was not at a phonological level, as presentation of the definition of one of a homophone pair (e.g. hare) did not prime later production of the homophone partner (e.g. hair). Thus any increased activation is not occurring in isolation at the phonological level. The second experiment used Welsh-English bilingual speakers (Monsell, Matthews and Miller, 1992). If the definition were presented in English, reaction times were faster only to production of the item in the same language, and not to production in the other language. They argue that if the effect were located within semantics, priming should occur across languages. These two sets of results indicate that the priming effect found in repetition of a target word lies not within either level of representation, but in the processing routines linking semantics to phonology. If this is the case, repetition of a word should aid aphasic people who have good enough semantics and good enough phonology so that the strengthening of the link between the two can impact upon the processes of word retrieval. Someone with significant damage to either semantics or phonology may not therefore benefit from this form of intervention.

4.2.1.2 Stimulating word production in aphasia

4.2.1.2.1 Phonological cues

Pease and Goodglass (1978) investigated the immediate effects on picture naming of a number of cues. They found the strongest effect from initial phoneme cues, even for severely aphasic people. The 'arousal power' of initial phoneme cues (Albyn Davis, 1993: 274) has guaranteed a significant amount of research devoted to the effect. Phonological cues have been shown to produce a dramatic effect in proper name anomia (e.g. Cohen, Bolgert, Timsit & Cherman, 1994). Avila et al (2001) report the results of cueing experiments with FR who had an anomic aphasia. They found a significant effect of phonological cues and a frequency effect in responsiveness to cues, with higher frequency words responding better to cues than lower frequency words. Lambon Ralph, et al (2000) report two single case studies of pure anomia and response to cues. For both GM and JS reading aloud whole words and initial phoneme cues had a significant effect on naming.

Monzell and colleagues' findings indicate that the normal cueing effect is located between semantics and phonology. It is not therefore surprising that people with pure anomia respond well to cues. In pure anomia both semantics and phonology are relatively well preserved. If the cues act on the link between these two levels they are likely to be most effective in such cases. In contrast, anyone with a deficit in either of these levels is likely to not respond to such cues. For example EST (see Chapter One, section 1.3.3.2.) had a phonological deficit and phonemic cues led only to further phonological approximations to the target.

One methodological issue which relates to research into the effects of cues and of facilitation concerns the use of a control condition. Most studies compare number of cues given with the number of times that cue was effective, giving a percentage value of the effectiveness of that cue. This fails to control for the fact that having extra time may also be effective and that, in the process of cueing, more time is made available for naming. It is important therefore that studies looking at this type of intervention include an extra time condition in which, on failing to name an item, the person is exposed to it for a further time period.

4.2.1.2.2 Phonological facilitation

A small number of influential studies have investigated the immediate and delayed effects of two types of phonological cue. The evidence from these studies is equivocal. Patterson, Purrell and Morton (1983) investigated the effects of initial phoneme cues and word repetition on immediate and delayed naming. They found immediate effects of both cue types but no lasting effects five minutes or thirty subsequent cueing events after administration of the cue.

Howard et al (1985a) looked systematically at the effects of phonological cues on word retrieval. They investigated phonemic cues and word repetition and found an immediate effect of the cues but no delayed effect 10 to 15 minutes after administration of the cue. This is in line with Patterson et al's findings. This was in contrast to the effects of semantic cues which led to lasting facilitation of word retrieval at 30 minutes and even at 24 hours after administration (Howard et al, 1985a).

Barry and McHattie (1991) looked at the facilitatory effects of a semantic cue and of word repetition. They found a delayed effect of the semantic cues at 20 minutes after administration, and a small but significant effect of repetition.

The studies differ in several important respects however. One significant difference is whether the picture of the target item is present or not. In Patterson et al's (1983) study the picture was present throughout. In Barry and McHattie's study it was not. A further difference concerns the time lag between provision of the immediate cue and the delayed naming test. How the tasks were presented also differs across studies. Howard et al (1985a) administered the two types of cues in separate blocks, whereas Barry and McHattie (1991) presented the two cue types within the same session. This may have influenced the outcome with some semantic effect generalising to the phonologically treated items.

The actual stimuli themselves constitute a further variable. In most studies these are pictures of common objects. They differ necessarily however in terms of the values of important psycholinguistic values, thus certain sets may be more difficult to name than others. A logical progression from this is that some items, although responding to initial phoneme cues, may be less robust in maintaining the increase in activation caused by the initial administration of the cue. If significant numbers of such items are present in a stimuli set, delayed effects of cues are unlikely. This is one possible explanation for the failure to find delayed effects of phonological cues.

4.2.1.2.3 Differences between semantic and phonological techniques

Two significant differences exist in the implementation of semantic versus phonological techniques. First the person may be exposed to the word form in the semantic technique (e.g. in being asked to point to the picture of the target word, or being asked about a specific attribute of the target such as "Does a cow eat grass?"). In the first phoneme cue condition, the person is not exposed to the word form in total, merely the first phoneme. Evidence for the importance of this factor comes from a single case study conducted by Le Dorze et al (1994). They compared semantic facilitation of spoken word production for a person with aphasia (RB) in two conditions. In one RB was exposed to the target word (the formal semantic technique), in the second he was not (the semantic technique). Three techniques were used in both conditions: written and

spoken word to picture matching and semantic judgements. The results showed a significant facilitatory advantage for the formal semantic technique, where the word form was presented to the person, over the semantic technique, where it was not.

The second important difference concerns the degree of choice involved in the phonological versus semantic conditions. In semantic techniques the person is required to make an active choice between a set of stimuli, for example in a word-to-picture matching task, or to determine whether a presented proposition is correct or not (e.g. 'Does a cow eat grass?'). In contrast phonological techniques involve no choice: the person is merely presented with the cue (part or whole word). One possible hypothesis explaining the advantage of semantic over phonological techniques concerns the theory of depth of processing usually invoked to explain certain phenomena found in research into memory (Craik and Tulving, 1975)¹⁷.

4.2.1.2.4 Orthographic facilitation

In contrast to the work done on the effects of phonological stimulation on spoken word production there is relatively little published work on the effects of orthographic stimuli. Bruce and Howard (1988) investigated the ability of people with aphasia to use grapheme to phoneme conversion routines to generate their own phonological cues from written letter selection. In a group of 20 people with aphasia they found no single participant with all the skills deemed necessary to achieve the desired result (selection of correct initial letter, ability to translate letters to sounds, and ability to benefit from initial phoneme cues). When naming was tested with a letter board of nine letters present none of the group showed an improvement in naming. Later work conducted by Howard and Harding (1998) however showed more positive results in a single case study of a woman with aphasia. Her spoken word finding improved significantly in the presence of the written alphabet. A further study (Basso et al, 2001) showed an advantage for orthographic cues over word repetition and reading aloud in two single case studies of people with aphasia and in a group of non-brain damaged participants (see section 4.2.3.3 below).

¹⁷ In Craik and Tulving's (1975) study participants were given words and questions about them. The questions involved "shallow" features, like font, "intermediate" features, like rhyming, and "deep" features involving meaning. They were then given an unexpected recall or recognition task. It was found that deeply encoded words were remembered better than shallowly encoded ones.

4.2.3 Therapy for spoken word production

4.2.3.1 *Design issues in aphasia therapy studies*

Over the past 30 years there has been a proliferation of studies of aphasia therapy reporting interventions for all of the known symptoms. It has rapidly become clear that there are significant issues regarding study design which need to be considered in setting up a therapeutic intervention experiment. Notably it is essential that therapy studies ensure that they control for the possible effects of the following variables as explanations for change in behaviour in the person with aphasia: spontaneous recovery; a generalised language therapy effect; a generalised effect of increased well-being as a result of the attention and time given by the researcher (see Howard, 1986, and Pring, 1986, for a discussion of these issues).

4.2.3.1.1 *Design and analysis*

There are now a number of published articles which denote possible therapy designs and their strengths and weaknesses (see Franklin, 1997; Nickels, 2002a). In order to control for the effects of spontaneous recovery a number of methods are generally used: one involves multiple baseline measures where a test is repeated many times prior to the intervention. If the scores are stable prior to therapy there are stronger grounds for claiming any change post-therapy to be due to the therapy and not the effect of another variable. One problem with this method is the risk of practice effects, demonstrated by Nickels (2002b) in her study of JAW's picture naming. This person with aphasia's performance on a picture naming task improved following attempts to name the picture, read the word aloud, and copy the written word after a delay. No feedback or correction was given and yet picture naming improved. A second consideration is time: if a test is lengthy repeating it several times may be contra-indicated. Often where repeated baselines are used the number of stimuli in the tested sets is relatively small.

A second possible method used to monitor any impact of spontaneous recovery involves the use of control *tasks* investigating performance in a language function which is not the focus of treatment. Few early studies used such tasks although this is becoming a more widespread practice. Change in these tasks would indicate a general improvement in language, and it would thus be difficult to attribute any change in treated tasks to the

intervention itself. Conversely where control tasks remain stable and treated language tasks improve there would be good grounds for claiming a specific therapy effect.

A third possibility is the use of untreated control *items* in therapy. It is important that the treated and untreated sets are matched for baseline performance and for key variables to eliminate the possibility of one set being easier than another. Many of the published treatment studies do not allow for statistical analysis of the results, by failing to use sets which are large enough to ensure adequate statistical power (e.g. Raymer et al, 1993). Another possibility is that both treated and untreated items improve as a result of therapy. In such an instance a specific therapy effect would be hard to claim unless evidence of no change in language control tasks were available.

Howard et al (1985b) included a naming control set in their study of semantic and phonological therapy. This set was presented along with the treatment set, but the person merely had to try to name the items and no input from the therapist was given. The rationale behind this decision is that any apparent therapy effect may be the result of the person trying to name the picture. Inclusion of a naming control allows for the comparison to be made. They found some effect of exposure to naming compared to unseen untreated items one week after therapy, but this effect had faded by six weeks post therapy. Many studies do not include such a set.

An ideal study therefore uses sufficient baseline testing with sets which are large enough to minimise any possible practice effect. How many baseline tests and how large the sets should be is a matter of debate but should take into account the experience undergone by the person with aphasia, the clinical applicability of the method, and statistical power given the available statistical tests. The study should include language control tasks which are unrelated to the language function being treated. Treated sets should be matched to untreated sets.

4.2.3.1.2 Relating breakdown in naming to therapy method

Within the tradition of cognitive neuropsychology attempts are made to isolate the source of a given individual's impairment and thus inform theories of normal processing. From this a clinical practice has emerged for speech and language therapists, who base language assessment upon a theory of language processing (e.g.

PALPA: Psycholinguistic Assessment of Language Processing in Aphasia: Kay, Lesser and Coltheart, 1992). It follows from this that the results of language assessment should inform the selection of language therapy. There are now a number of studies which seek to identify the source of the impairment and relate this to the resulting therapy selection and results (e.g. Nettleton and Lesser, 1991; Nickels and Best, 1996b). There has been an ongoing debate about the capability of cognitive neuropsychology to provide guidance for therapy (see e.g. Caramazza, 1989; Caramazza and Hillis, 1993; Best and Nickels, 2000). Although some clear-cut results have emerged (e.g. Nettleton and Lesser, 1991) it remains unclear which therapies work for which participants with aphasia.

Nettleton and Lesser (1991) investigated whether therapy based on a hypothesised level of breakdown was effective compared with a therapy which was inappropriate given the person's level of breakdown. They looked at three levels of breakdown with two people with aphasia in each category: semantic, phonological lexicon, and phonemic buffer. Four people were given appropriate therapy for their profile. The two people with a semantic deficit were given semantic therapy. One showed an increase in treated items, and the second showed a change in error pattern towards more closely related semantic paraphasias. The two people with a phonological lexicon deficit received phonological lexicon therapy, which consisted of word repetition, naming with progressive cues and rhyme judgements. Both improved, with one showing a generalised effect to untreated items. Finally the two people with an output buffer deficit were given semantic therapy, which was not effective for either of them. There are problems in the interpretation of the data however. One of the people with a semantic deficit was possibly still undergoing spontaneous recovery and the results may be due to that and not to the therapy. In addition there were no language control tasks and thus the generalisation effect seen in one of the people with phonological lexicon damage may be a non-specific effect of therapy. Although there are some problems with the design of this study, the results are nevertheless encouraging for the belief that psycholinguistic theory can influence therapy decision-making.

It is by no means the case that semantic therapy is effective only for people with a semantic deficit however. Marshall et al (1990) report the results of therapy involving word to picture matching tasks for three people with aphasia. This small group had semantic deficits ranging from mild to moderate, yet the therapy was effective for all of

the participants regardless of severity of the problem. Thus it may be that a therapy involving a semantic task can be effective for a range of semantic involvement. An alternative explanation is that the so-called semantic task activates both semantic and phonological representations. The way in which any given individual benefits from this composite therapy may differ, depending upon the source of their word finding deficit.

Best and Nickels (2000) state that a number of factors need to be addressed to improve research in this area: more fully specified theoretical models, simpler therapy tasks which allow a transparent evaluation of the putative language processes involved, and fuller reporting of the linguistic profile of individuals.

4.2.3.1.3 Generalisation to untreated items

A further issue of crucial importance for the usefulness of therapy techniques trialled in research concerns the generalisation of therapy effects. The term generalisation is used to denote two distinct issues: generalisation of therapy effects from the treated items to the untreated control items, and generalisation of spoken word production from picture naming tasks to other speech tasks. For the purposes of clarity the latter issue will be termed 'carryover' throughout this report. In order to investigate the former, matched sets of untreated and treated items are generally used.

The main findings so far are not encouraging with few studies showing significant change in untreated items. Of those that have the majority are studies implementing semantic therapy. Marshall et al (1990) used word to picture matching to improve spoken word production. They found generalisation to untreated items in one of the three participants (FW). Nickels and Best (1996b) used semantic judgement tasks with AER and TRC and showed generalisation to untreated controls for some of the subsets. In contrast therapy highlighting or stimulating the phonological form of the target has on the whole been less successful in achieving generalisation. Greenwald et al (1995) used a phonological cueing hierarchy, and word repetition to treat word finding in SS, and failed to find evidence for generalisation. By contrast however, Robson, Marshall, Pring and Chiat (1998) found an improvement both in untreated items that were phonologically related to the target and in those that were not. This therapy differed from other phonological therapies as it involved reflection upon the word form (number

of syllables and first phoneme) and may therefore tap into different processing routines from more strictly phonological therapies such as word repetition or phonemic cueing.

4.2.3.1.4 Carryover to other speech situations

It is important that any therapy which aims to improve word-finding in aphasia can demonstrate an effect on the real life communication of that person. Carryover to speech tasks beyond picture naming has proved difficult to measure. It is true that few studies looking at the effects of treatment for word-finding deficits have investigated whether carryover occurs. It is possible that such carryover is present but that measures used to detect it have not been appropriate. It is certainly the case that certain methodological issues influence the administration of such measurements.

There is moreover a distinction to be drawn here between connected speech tasks such as story telling, which arguably lack ecological validity, and functional carryover. The latter can only be measured by samples of real conversation, and observation of the person with aphasia's communicative behaviour in real contexts. In most studies the former tasks are used, with few studies reporting conversation data.

The connected speech tasks which have been used include picture description and word definition (Davis and Pring, 1991), connected speech samples (e.g. Spencer, Doyle, McNeil et al, 2000); request for procedural information, request for personal information, picture sequence description (Coelho, McHugh and Boyle, 2000) and asking the person to give a verbal description (Best et al, 1997). Several of these studies used the measurement of what are termed 'correct information units' or CIU's (Nicholas and Brookshire, 1993) to analyse the connected speech samples. Single words are counted as CIU's and are defined as "accurate relevant and informative relative to the eliciting stimulus" (Nicholas and Brookshire, 1993: 340). Recently Franklin et al (2002) used a sentence generation task to tap into any carryover effects of therapy for MB. Carryover of production of items treated in the therapy sessions was tested by presenting MB with pictures showing the same targets in action pictures.

The methods outlined above suffer from the same limitation as picture naming in that they do not sample everyday communication directly. A range of assessments have been developed which attempt to do just that. As Lomas et al (1989) state, in order to assess

functional communication it is necessary to directly observe or obtain a report of direct observation of the person with aphasia engaged in actual communication situations. This is done by direct observation, role play, or reports of communication by carers, and often involves rating the person's ability on a numerical scale in a range of speech situations. A number of such assessments now exist (e.g. Communicative Effectiveness Index: Lomas et al, 1989; Communicative Abilities in Daily Living: Holland, 1980, Amsterdam-Nijmegen Everyday Language Test: Blomert et al, 1987).

Conversation analysis provides a more in-depth exploration of the interaction between the person with aphasia and their conversation partner. Detailed analysis of the conversations between a person with aphasia and their main speaking partner/s can provide possible targets for therapeutic change, often targeting the speaking partner's conversational behaviours (see e.g. Booth and Perkins 1999). A clinical consideration which limits the usefulness of conversation analysis is the amount of skill, training and time required to transcribe conversation samples competently.

Possibly for those reasons very few studies have investigated conversation. Robson et al (1998) report improvement in conversation following phonological therapy for the person with aphasia GF, but provide only a short sample of conversation with no analysis. This area remains in need of development as, although improvement in tasks such as picture description are noteworthy, the same criticism that has been levelled at picture naming as an outcome measure may be levelled at the above tasks: they are not normal speaking behaviours, and thus may not be valid methods of addressing the issue.

4.2.3.1.5 Maintenance of therapy effects

A final consideration is the issue of maintenance of any therapy effects. This is analysed by reassessment of therapy and control items at a time lag after the end of the therapy intervention. Most studies have tested at one to two months post therapy (e.g. Fink et al, 2002; Howard et al, 1985b; Marshall et al, 1998). The results are on the whole encouraging with most positive therapy outcomes maintaining at least to some degree after a time lag. Investigations of the longer term effects of therapy are rare, although two studies have looked at longer lags. Pring et al (1990) followed up the group of people studied in Marshall et al (1990) and found therapy effects maintained one year

after therapy had ceased. Miceli et al (1994) found continued positive effects on treated items for one person 17 months after therapy had ended.

4.2.3.2 Semantic therapy

A brief summary of semantic therapy is provided here. Semantic therapy is the term used to describe an intervention which intends to strengthen or activate semantic representations. Therapy tasks involve semantic judgements, and include word to picture matching, attribute judgements, and semantic categorisation. Tasks differ as to whether they also involve production of the target word. Intervention for word finding difficulties using semantic techniques has been shown to be effective in a number of studies (e.g. Howard, Patterson, Franklin, Orchard-Lisle and Morton, 1985b; Marshall, Pound, White-Thomson and Pring, 1990).

Howard et al (1985a) found that semantic facilitation has relatively long lasting effects, whereas phonological facilitation has short lasting effects (see section 4.2.1.2.2 above). In an extension of this work Howard et al (1985b) compared semantic and phonological facilitation in a treatment study. People had daily therapy for one or two weeks. The advantage of semantic over phonological techniques disappeared with repeated administration of the cues over time, although there was some evidence of greater generalisation to untreated items with the semantic approach than with the phonological approach.

The data in Howard et al (1985b) are difficult to interpret as only aggregated post therapy scores for the group of people with aphasia are given. In a recent re-analysis of the data Howard (2000) reports that the claim for an advantage of semantic over phonological therapy may have been premature: "The differences between the two techniques are minimal: the only significant difference is for better performance with semantic naming controls than phonological naming controls one week after therapy" (Howard, 2000: 85).

It may then be that phonological therapy is as effective as semantic. Key factors to consider in evaluating the following therapy studies are: what the actual therapy consisted of and, crucially, whether there was a semantic element to the therapy; whether the therapy had an effect on only the treated items or whether the effect

generalised to untreated items; and the nature of the individual with aphasia's word finding deficit.

4.2.3.3 Phonological therapy

The term phonological therapy is used to denote interventions which intend to activate the phonological form. The person will usually be presented with a picture to name and be offered a facilitatory cue when unable to find the word. There is also in all likelihood semantic activation from the processing of the picture, in addition to the intended phonological activation from the cue. Tasks used include initial phoneme cues, progressive phonemic cues, rhyme cues, reading the word aloud, and repetition. A review of these therapy studies is offered here as this approach was used with the three participants reported in this thesis in the first phase of therapy.

Recently a number of studies have shown that phonological approaches to the treatment of word-finding difficulties can be effective (e.g. Davis and Pring, 1991; Raymer, Thompson, Jacobs and Le Grand, 1993; Hillis and Caramazza, 1994; Greenwald et al, 1995; Miceli, Amitrano, Capasso and Caramazza, 1996; Robson et al, 1998; Fink, Brecher, Schwartz and Robey, 2002). The studies selected here are ones in which there is an adequate account of each person with aphasia's word finding deficit, and in which necessary steps are taken to control for spontaneous recovery, and general language stimulation effects.

In addition to the studies which report the effects of clinician-administered and computer-administered cues, a number of studies exist which have sought to develop self-cueing techniques by promoting the translation of letters into sounds (Bruce and Howard, 1987; Nickels, 1992; Best et al, 1997). These will be discussed in section 4.2.3.3.8.

4.2.3.3.1 Davis and Pring (1991)

Davis and Pring (1991) investigated the effects of three forms of intervention with seven people with aphasia, all of whom were beyond the phase of spontaneous recovery. The three forms of intervention included i) word to picture matching with related distractors with reading aloud or word repetition; ii) as for i) but with unrelated

distractors; iii) the correct name of the target was presented and the person was asked to repeat the name. One set of 30 stimuli were treated in each condition. Ten sessions of therapy were implemented and each item was treated twice in each session. The results showed that all three treatment sets improved significantly after the intervention. A set of untreated unrelated distractors also improved although a set of untreated related distractors did not. The authors also investigated the generalisation of treatment effects to other speech tasks through picture description and spoken word definition, where significant improvement was also found.

Best and Nickels (2000) call for more clearly defined therapy tasks. In Davis and Pring's (1991) study two of the therapy conditions included both semantic and phonological tasks. Of interest here is that the repetition condition fared as well as the two conditions which had both a semantic and a phonological component. The data from the two untreated sets were an enigma and remain so. Further studies comparing semantic relationships between treated and untreated items are warranted.

4.2.3.3.2 Raymer et al (1993)

Raymer et al (1993) describe four participants with aphasia in terms of the single word processing model taken from Ellis and Young (1988). One person (CG) presented with a deficit affecting the phonological output lexicon primarily, whilst the other three (RJ, MR and RE) presented with deficits affecting lexical-semantic processes and possible further phonological output lexicon deficits. The treated items were presented for naming. If the person failed to name a picture three cue types in progression were administered until the person named the word, or failed after the third cue. The cues were spoken presentation of a rhyme word, initial phoneme cue, and spoken presentation of the target word.

The study used a multiple baseline design with repeated probe measurements of naming of the target and control words throughout the intervention phase. The experimental stimuli consisted of two sets of 30 words, each divided into three further sets: a set of 10 treatment items, a set of 10 words which rhymed with the treatment items and a set of 10 words which had a semantic relationship with the treatment items. All sets were probed throughout to measure any generalisation of therapy effects to untreated related (rhyming or semantic) items. The small number of items, and unstable baselines make

statistical analysis of such results difficult although tests of trend could have been used. The authors merely report the percentage correct in the various sets throughout the study, combining the results of the rhyming and semantic sets. The results appear to show that the treatment was mildly effective in increasing naming success for the treated items, and some weak evidence for generalisation to other language functions, notably reading aloud single words for two of the participants (CG and RE). Assessment of maintenance of therapy effects was attempted at two months post intervention and showed some evidence for maintenance in the three participants who completed this part of the study (CG, RJ and RE). Again, no statistical analysis is provided.

Although there are some design and methodology problems with this study it was one of the first to show that phonological techniques could be effective. What is not clear is which part of the therapeutic intervention led to the increases, such as they were, in naming success. Three methods were used and any one of these could have been crucial. By using all three it is not possible to differentiate the independent effects.

4.2.3.3 Hillis and Caramazza (1994)

In a series of studies Hillis and Caramazza (1994) explored the relationship between therapy and level of deficit. In study three of this series they worked with two people with aphasia. JJ had a lexical semantic deficit, making errors on word to picture matching tasks and producing semantic errors in both written and spoken picture naming. HW by contrast made no errors on word to picture matching or written picture naming, yet made semantic errors in spoken picture naming and reading aloud. Hillis and Caramazza (1994) argue that HW's deficit lies in access to or within the phonological output lexicon.

The participants received the same two forms of therapy, one semantic and one phonological. Items were treated in all sessions with either written word to picture matching, or with reading the word aloud with phonemic cues to facilitate production. Items were randomly assigned to treatment conditions on a session by session basis. The facilitatory effects of the two interventions were measured in each succeeding session. By the end of the therapy however, any given item would in all likelihood have been treated by both techniques. The results are extremely neat. The assessments in

each session showed that JJ benefited significantly from the semantic task but not from the phonological task, and vice versa for HW.

JJ and HW both present with a clear-cut profile on psycholinguistic testing. The facilitation methods used suggest that a semantic deficit may respond to a semantic technique, and that a lexical-phonological level deficit may respond to a phonological technique. It may be more difficult to find an appropriate therapy however when there are multiple levels of deficit.

4.2.3.3.4 Greenwald et al (1995)

Greenwald et al (1995) investigated the effects of what they term a 'phonological cueing hierarchy' with two people with aphasia, SS and MR. Both presented with impaired semantic processing affecting input and output modalities. Their spoken word production was severely impaired, and both made semantic errors on spoken and written picture naming. The authors administered this form of intervention with a subsequent semantic cueing hierarchy therapy in order to target first the lexical retrieval impairment in the two participants, then the lexical semantic impairment. Given the nature of the deficits and the therapy implemented it is unclear how this study can contribute to the debate concerning the relationship between word-finding breakdown and therapy effect.

The phonological treatment consisted of naming a word from a spoken auditory definition. If unable to do so the person was asked to provide the first sound, and if unable to provide that they were told "It begins with x". If this failed to elicit the target word they were then given the first two sounds, with the same procedure, then finally the target word. The treatment items were 40 pictureable nouns split into two sets. Within each set, half were trained and half untrained. Items were paired in the trained and untrained sets for semantic category. Thus 10 items in each set received the training. Each item was trained two to four times per session in sessions lasting up to one hour, and treatment took place for up to 20 sessions.

The authors claim significant results for treated sets for SS and MR, and some effect of generalisation to untreated items for both participants. Although there may be positive outcomes for this form of therapy for people with marked semantic impairment and severe anomia, there are two obvious difficulties with the study which make the results

hard to interpret. First, the therapy incorporates semantic and phonological cues and thus it is impossible to distinguish the two effects. Second, there are very few items in the various sets. A final issue about this study concerns the amount of therapy implemented, which was up to 20 sessions, for such a small set of words.

4.2.3.2.5 Miceli, Amitrano, Capasso and Caramazza (1996)

Miceli, Amitrano, Capasso and Caramazza (1996) report the results of a further phonological therapy study. They locate the rationale for the choice of therapy method in an analysis of the hypothesised source of the word finding deficit. The authors state that in a model proposing mappings from semantic to phonological output lexicon entries, therapy should achieve item-specific effects as the mappings are achieved in a one to one relationship. Results of language tests for both participants are shown in Table 4.1 for ease of comparison.

**TABLE 4.1 Results of relevant language tests for
RBO and GMA (Miceli et al, 1996)**

Task	RBO		GMA	
Picture naming	161/300	0.54	396/500	0.79
<i>Semantic tasks</i>				
Spoken word to picture matching	40/40	1.00	39/40	0.98
Written word to picture matching	37/40	0.94	39/40	0.98
<i>Phonological tasks</i>				
Reading aloud words	69/92	0.75	184/192	0.96
Repetition words	42/45	0.93	45/45	1.00
Reading aloud non-words	29/45	0.64	41/45	0.92
Repetition non-words	22/36	0.61	35/36	0.97

RBO was a 38 year old woman seen 12 months post onset who presented with good lexical semantic processing, damaged but partially functional sub-lexical transcoding mechanisms, and deficits in word repetition, reading words aloud, and in picture naming. The authors propose that her deficit lies within the phonological output lexicon. The experimental stimuli consisted of 90 words which RBO had failed to name at three consecutive test times and for which she showed good comprehension. These were split into three sets of 30 items. Intervention for the first set involved the presentation of the written word for RBO to read aloud, for the second set presentation of the spoken word

for RBO to repeat. The third set acted as untreated controls¹⁸. The two interventions were carried out on the respective sets at a time lag between the start of the two interventions of three days. Intervention involved five consecutive sessions of one hour each for each treatment set. Follow-up was carried out 25 days after completion of the first treatment and 17 days after the second.

The results of the interventions were analysed using McNemar's tests. Both interventions led to significantly improved picture naming, and there was no generalisation of therapy effects to untreated items, in line with the hypothesis proposed regarding RBO's deficit within the phonological output lexicon. There was however an increase in performance in untreated items between the two therapies. As Nickels and Best (1996a) point out this could reflect a delayed effect of therapy. Therapy effects maintained at follow-up 25 and 17 days after the conclusion of the two interventions.

The second person discussed in this paper is GMA who presented with a similar pattern of deficit to that seen in RBO, but to a milder degree. GMA had good lexical semantic processing, a mild deficit in non-word transcoding, good repetition and reading aloud and a marked picture naming deficit. The authors propose that, like RBO, GMA had a deficit affecting the phonological output lexicon. As for RBO therapy items were selected from those targets GMA never achieved consistently in three consecutive naming attempts but for which he showed good comprehension. The 80 items were split into four sets, three of which received intervention for seven consecutive days each in one-hour sessions. Each item was treated 10 times in each session, "and errors were corrected as many times as necessary until the correct response was produced" (Miceli et al, 1996: 164). Treatments occurred consecutively with a seven-day break between each period of treatment. The three conditions involved: i) presentation of both the target picture and the written word; ii) written word presented alone; iii) the target picture was presented and if GMA failed to name it increasing amounts of the phonology of the target were spoken aloud by the researcher (initial sound, initial syllable, first two syllables etc.) until the correct response was provided.

The results of this study showed a significant change in each treatment set as a result of the treatment. All three interventions were effective as shown by McNemar's tests, and

¹⁸ It is unclear from the case description of the intervention whether the target picture was present or not. This is crucial in determining the effects of the interventions and in replicating the study.

no one intervention was significantly more effective than any other¹⁹. The untreated control set remained at baseline, and the sets not being treated during any particular treatment phase also remained stable, indicating that no generalisation of therapy effects occurred. The authors were able to follow up GMA over the next 17 months and found good maintenance of the therapy effects, suggesting a robust effect of the interventions.

This well-designed and thoughtful study shows that therapy targeting the phonological form of the words, in two people whose word finding deficit is thought to derive from damage to the phonological output lexicon, can be effective and can have a lasting effect. The therapy effect was confined to the treated items. What is not clear is the specific nature of the damage to the output lexicon. The severity but also the pattern of impairment differs across the two participants.

4.2.3.2.6 Robson *et al* (1998)

The study considered here was carried out by Robson *et al* (1998) with GF who presented with jargon aphasia. GF had good semantic processing when accessed from spoken input, a severe picture naming deficit, but better access to phonology from written and spoken input in reading aloud and repeating words. A summary of GF's profile is shown in Table 4.2. GF had an imageability effect in word repetition and most of her naming errors consisted of failures to respond. The authors noted in the assessment phase that GF responded well to phonological cues, and showed some knowledge of the syllabic structure of items she could not name and they therefore selected a phonological approach for the therapy.

Therapy consisted of 40 sessions of 20 minutes per session carried out over a six month period. A set of 24 items were selected as the treatment set, and there were also 24 phonologically related untreated items, and 24 phonologically unrelated untreated items. The therapy encouraged GF to think about the syllabic structure and the first phoneme of the target, targeted in the first sessions in separate tasks and in later sessions in the same task. In the earlier sessions GF was not asked to name the picture, but the target was presented to her for repetition. In later sessions she was encouraged to reflect on the phonology and then to attempt the target name.

¹⁹ Evidence for this claim comes from consideration of the numbers of items correct before and after therapy and not from the authors' analysis of this possibility.

**TABLE 4.2 Results of relevant language tests for
GF (Robson et al, 1998)**

Task*	GF	
Picture naming	1/40	0.03
<i>Semantic tasks</i>		
Spoken word to picture matching	39/40	0.97
Written word to picture matching	28/40	0.70
<i>Phonological tasks</i>		
Reading aloud words	10/40	0.25
Repetition words	19/40	0.48
Reading aloud non-words	-	-
Repetition non-words	3/30	0.10

*All tests used are from PALPA.

The results provide pre and post therapy scores in the three sets and also scores at follow-up eight weeks after the end of therapy. GF made significant gains in picture naming in the set of treated items, the set of phonologically related controls and the set of unrelated controls, as shown by McNemar's tests. These gains were maintained at follow-up. There was no change in other language tasks used as controls (e.g. Pyramids and Palm Trees). The authors offer a positive qualitative account of changes in GF's conversation after therapy but offer only short samples to support the claims.

In this study a substantial amount of therapy is devoted to a relatively small set of items. Given that the therapy effect extended to control items this amount may be justified. This study is confusing in that the therapy was effective, in the context of the person being unable to carry out the therapy tasks (identification of first phoneme, and number of syllables) when these were tested in isolation. It is therefore extremely unclear what made the difference to picture naming.

4.2.3.2.7 Fink et al (2002)

Fink et al (2002) provide further data concerning the effectiveness of phonological cues for naming, but in this case delivered via computer. They set out ostensibly to compare two conditions: one where the therapist monitors and guides the therapy via computer, and one where the person with aphasia guides their own therapy. With three people in each group and no attempt made to match participants across groups this aspect of the study appears so methodologically flawed that no conclusions can be drawn from it.

Notwithstanding this problem, the study does provide further data on who can benefit from this type of therapy.

The therapy consisted of picture naming with cues to help if the person could not name a picture. The cues were hierarchical and consisted of initial phoneme or letter, sentence completion, and whole word repetition or read aloud. After this series of cues the person then repeated the word aloud three times. For each individual the optimum level and modality of cueing was established, and this served to guide the therapy procedure. For example one individual might be better with written than spoken cues and at a whole word level rather than first letter or sentence completion.

Those items which an individual scored worst on in four administrations of the Philadelphia Naming Test were allocated to the two treatment sets ($n = 20$ per set). A crossover design was used with one set being treated and the other remaining untreated then vice versa. Maintenance was assessed four weeks after the end of therapy.

The results showed a positive item-specific effect of therapy for five of the six participants. One participant (BM) showed little response to the therapy. BM was the only person to present with a significant semantic deficit, and he was totally unable to read aloud²⁰. All of the other participants had no or little semantic involvement. All participants had good transcoding skills in at least one of reading aloud and repetition. Two of the five also showed some effect of generalisation to untreated items. There was no attempt to measure carryover to everyday speech situations.

It appears that the integrity of the semantic system and the phonological output lexicon are important in order for a phonological therapy such as this to be effective. All participants for whom therapy worked had good semantic processing and preserved phonological representations. The one person for whom therapy did not work had impaired semantics and a total deficit in one transcoding task.

²⁰ This latter point may not be important as the therapy offered either phonological or orthographic cues and BM's therapy, one assumes, used the former.

4.2.3.3.8 *Self-cueing phonemic cues*

In all of the studies reported above the therapist or computer administers the relevant facilitator, leaving the person with aphasia a relatively passive recipient. Bruce and Howard (1987) wondered whether people with aphasia could learn to provide their own cues. They investigated people with Broca's aphasia's ability to generate phonemic cues using an electronic aid. They identified five people with aphasia who were able to identify the first letter of words they were unable to name, and were also able to benefit from phonemic cues. The computerised aid acted as the missing link in the process by converting a letter button pressed by the person into an auditory sound cue. They found significant differences across the five participants, but an overall significant effect of the aid. The treatment effect did not however generalise to untreated items, indicating that the participants had not learnt to use the aid but had developed an item-specific knowledge.

Nickels (1992) worked with TC who had a semantic deficit and a deficit affecting access to the phonological output lexicon from semantics. Despite a severe deficit in oral naming TC was able to generate some written names and read these aloud. Nickels (1992) used the relay method developed by De Partz (1986) to try to develop spelling-sound correspondences in TC. This method involved training the person to associate a known word (e.g. a family member's name) with a letter, thereby eventually enabling the person to segment off the first sound and associate this with the letter. TC was able to make associative links between words and letters, and to segment off the first sound. He was unable to blend a sequence of sounds together however and made no progress on this aspect throughout the therapy study. Nevertheless he made significant improvements in reading aloud and naming.

Best et al (1997) used the computerised aid with JOW who had a deficit in accessing phonological output lexicon representations from relatively intact semantics. JOW showed some knowledge of written word forms, and was able to respond to phonemic cues. Therapy encouraged JOW to think of the written form and press the appropriate letter key on the aid. The results showed an improvement in naming treated items and untreated control items. Assessment post therapy compared naming with the aid and naming without the aid and found no difference between the two conditions indicating that the improvement in naming was not dependent upon use of the aid, but that JOW

had perhaps internalised the component parts of the therapy to enable better word retrieval.

4.2.3.4 Orthographic therapy

There remain relatively few studies investigating the effects on spoken word retrieval of the implementation of orthographic cues. Jokel and Rochon (1996) compared the effects of having the written word present, repeating the word, and sentence completion on the picture naming of an individual with aphasia (PD) who presented with a pure anomia. She had a deficit in accessing the phonological output lexicon from semantics. She made no errors on tests of semantic input processing, reading words aloud or repetition, yet had a significant deficit in naming pictures. They found a significant advantage for the written word over the other two cue types.

Basso et al (2001) compared the effects of three forms of intervention on non-aphasic participants' learning of novel words and on the picture naming of two people with aphasia. Little detail of the language impairment of the people with aphasia is given except to say that both were anomic, and both made mainly no responses or failed to name. The three methods trialled were: reading the written word aloud, repeating the target name, and the provision of an orthographic cue. For the non-aphasic control participants the orthographic cue led to better learning of new words when compared to the other two methods. The methods used were the same for the people with aphasia, except that for the people with aphasia the words trained were real words that they had been unable to name in a picture-naming test. These words were split into four sets, three treatment sets and one untreated control set. Treatment involved presentation of the target picture and then provision of the cue. The results showed a significant advantage for the orthographic cues over the other two forms of intervention and these did not differ from each other. Unfortunately inadequate detail about the two people with aphasia's language function is given to allow any hypotheses regarding the replication of the results with other people with aphasia. Nevertheless the results are encouraging for the implementation of orthographic cues as a therapy technique.

4.2.4 Therapies promoting communicative use of words

In the therapies described above the emphasis is on training a set of target words. There exists within the field of aphasia therapy a set of approaches which encourage communication of the concept or idea rather than successful production of the target word. These approaches, which usually come under the banner heading of 'functional' therapy, include a myriad of methods and theoretical stances. They have in common that they do not address the language impairment directly. These methods seek to bypass the language deficit by encouraging the person with aphasia's use of other modalities such as gesture or drawing. The best known of these is PACE (Promoting Aphasics' Communicative Effectiveness: Davis and Wilcox, 1985). Such approaches encourage the person to use whatever communication means are at their disposal to achieve effective communication of the target message. Alternatively therapy may focus instead on the conversation partner and the interaction between the partner and the person with aphasia. Therapy in such cases consists of the modification of the partner's communication behaviours in targeted contexts (e.g. Kagan, 1998).

One problem with this set of methods is that of demonstrating efficacy (Holland, 1991). Quantitative measurement of the range of communicative behaviours involved in such therapy is problematic. This is highlighted by some of the choices of outcome measures used by researchers in the field: Pulvermuller and Roth (1991) used the Token Test to assess the effect of a communicative therapy developed from PACE, and Aten, Caliguiri and Holland (1982) used the CADL (Communicative Activities of Daily Living: Sarno, 1969) and the PICA (Porch Index of Communicative Ability: Porch, 1967) to investigate the effects of a functional communication therapy involving role plays of everyday situations. Unlike the studies described in section 4.2.3.3, in these studies there is no set of treated items to test prior to and after therapy, so researchers are reliant often on published tests which may not be valid measures of the behaviour in question, or if they are valid measures, may not be sufficiently sensitive to identify small changes in the behaviour undergoing treatment.

Applying the strong methodology of the studies developed in the field of cognitive rehabilitation (outlined in section 4.2.3.1) would provide a more effective means of evaluation of such approaches. One study has attempted to draw the two approaches

together. Springer et al (1991) compared two therapy methods: a traditional PACE task, and what they term a modified PACE task. In the former the person was given a card with a picture representing a known object. They then communicated that concept to the therapist by whatever means they found most effective. In the modified PACE condition the person carried out 'language systematic training' as well as the communication of the message. They were given 22 pictures, from which they had to sort a particular semantic subset, such as tools, toys, fruit, the names of the subsets being written on cards. The therapist also had the 22 pictures and the written category name. A screen separated the person with aphasia and the therapist. The person with aphasia then communicated to the therapist the decisions made regarding category membership of each item. The therapist gave feedback on the accuracy of the decision. Each person with aphasia received four periods of treatment lasting five days each.

The study involved four people with aphasia, three of whom presented with a lexical semantic deficit, and one of whom presented with an apraxia of speech. The authors predicted that the former three would benefit from the semantic therapy input of the modified approach while the fourth person would not. Outcomes were measured in terms of the accuracy of production of the target names, and by means of the traditional PACE evaluation of communication. The results showed a significant effect of the modified PACE method on word retrieval and on communication of the target, for the three people with lexical semantic deficits. The person with apraxia of speech did not benefit from the modified PACE approach but did benefit from the traditional PACE therapy.

The results demonstrate that a cognitive approach to therapy, which isolates an impaired language function and directs therapy to the remediation of that function, can be paired with a therapy which encourages communication of the message. The study lacks baselines and adequate controls, but nevertheless offers a method of extending the traditional cognitive therapy involving cueing picture naming to a broader application.

4.2.5 Conversation-based therapy

Approaches which target conversational behaviours directly fall into two categories. Functional pragmatic approaches such as Holland's (1991) conversational coaching, and treatments which are drawn from conversation analysis (CA). The former target specific

behaviours in the real life setting and promote change in both the person with aphasia and their speaking partner. The second approach (CA) identifies patterns within the two speakers' conversations and targets specific behaviours for change. Again, quantifying any resulting change is problematical as little data exists concerning the reliability of quantification of conversation behaviours across test times. In addition proponents of CA insist that conversation cannot be quantified and that qualitative analysis is the only appropriate method of analysis. One problem with this set of approaches is that of demonstrating efficacy. Quantitative measurement of the range of communicative behaviours involved in such therapy is problematic due to lack of agreement regarding what behaviours are significant and subjective ratings of participants' ability being the usual method of assessment

4.3 SUMMARY AND IMPLICATIONS

This review of the current state of research into the effectiveness of therapy for word finding deficits in aphasia has identified some key issues which will be addressed in the work detailed in this report. A number of studies have shown that therapy for word-finding deficits which targets the phonological form of the word can be effective. It is unclear at present which people with aphasia, in particular which forms of spoken word production deficits, will benefit from this form of intervention. There is limited but growing evidence that orthographic cues applied in therapy can lead to gains in picture naming. Again, it is unclear which forms of spoken word production deficit will respond to this form of intervention, and the degree to which function in orthographic lexical and sub-lexical processing mechanisms needs to be retained in order for the intervention to be effective. The processes involved in recovery and the principles of intervention are beginning to be better understood although models of recovery are still in their infancy. The implementation of a therapy should be based on an understanding of brain recovery processes and the means by which brain function is regained through cortical plasticity. According to Robertson and Murre (1999) maximal recovery will depend upon i) principled stimulation based on theoretical models of mental processing, ii) significant and adequate amounts of repeated administration of the stimulation, iii) the attention of the person being directed to the stimuli, and iv) adequate arousal levels in the person concerned. Without all four of these criteria guided recovery will not be effective. Early reports of gains in picture naming after semantic facilitation, which maintained over time, have led to this method being preferred over phonological

approaches. However the element of choice distinguishes the two methods. It could be the case that a choice of phonological cues will lead to longer lasting effects of intervention from phonological methods. There are very few reports of therapy which use the strong methodology of the cognitive approach to intervention but aim to maximise communicative use of the stimuli. Further investigations of this type are warranted. Although there are many reports of therapy for word-finding deficits which report gains in picture naming, there are very few attempts to measure the impact of the therapy on speech situations other than picture naming. Those studies that have looked at this issue have used tasks such as picture description, which, it may be argued, lack ecological validity. It is important that methods to investigate whether picture naming gains transfer to everyday conversation. To this end specific quantitative measures are required.

CHAPTER FIVE: DESIGN AND METHODOLOGY

5.0 INTRODUCTION

In this chapter the overall design of the study conducted for this thesis is described. This includes an overview of the study, details of the therapy methods and design, and details of the outcome measures used in the study. A number of participants with aphasia took part in the overall study, details of which are reported in Best et al (2001), Hickin et al (2002) and Herbert et al (2003). The data from three participants are included in this report.

5.1 OVERVIEW OF THE STUDY

5.11 Participants

This research concerns the nature of the word finding deficit in aphasia, and the effects of therapy thereon. It was considered important to restrict the study to people who had sustained a single left hemisphere vascular lesion. Restriction to this population, who form the bulk of those presenting with aphasia, ensures that the results can be validly compared across participants within the study, and findings generalised to other members of the population outside the study. It also eliminates to a large extent the possibility of the language impairment emerging from a generalised cognitive impairment such as might be the case with other aetiologies such as traumatic brain injury.

As this research investigated the effects of therapy it was important to eliminate the possibility that improvement found after therapy was not due to the therapy itself but was an effect of generalised improvement found in the period of spontaneous recovery. For this reason only participants who were at least one year post onset of stroke were included.

The research focused on word finding difficulties and it was therefore important to identify people who had aphasia with significant word finding deficits, and whose speech was intelligible. This meant the exclusion of those with dyspraxia of speech or with dysarthria, and ensured that any errors produced in spoken word production could

be analysed in terms of a disruption to the lexical production system, and were not due to an impairment to motor speech processes.

A number of other criteria were implemented, the remainder of which are self-explanatory. The full set of criteria are: the participants must be at least one year post onset of a single left hemisphere CVA; present with aphasia with significant word finding difficulties; must have intelligible speech with no significant dyspraxia of speech or dysarthria; have no significant hearing or visual impairment; and have no other significant neurological history.

The three participants included in this thesis met all the above criteria. In addition the following criteria were applied to the selection of participants for this thesis. At least one participant should present with a word finding deficit implicating the mapping processes between semantics and the phonological output lexicon. This is because it was hypothesised that one of the two therapies used here – the lexical therapy – would be most effective for this form of word finding difficulty. A second participant should present with a deficit in word finding affecting processing in semantics and/or phonology. The inclusion of these two forms of deficit would allow the comparison of the effectiveness of the therapy to be made across the two deficit types. A third participant was included who presented with a deficit in mapping between semantics and the phonological output lexicon, but whose overall profile was more complex than that of the first person to be included with this form of word finding difficulty, with impairments in particular in orthographic to phonological processes. As the lexical therapy used phonological and orthographic cues, this would allow the comparison of the effectiveness of the two cue types across the two domains, phonological and orthographic, with overall processing within those two domains, within each participant.

5.1.2 Design

The study described here consisted of four stages, each lasting approximately eight weeks. This design is outlined in Table 5.1. Assessments One and Two occurred before the start of therapy Phase One and constituted the baseline. Assessment Three occurred immediately after the end of therapy Phase One, Assessment Four immediately after the end of therapy Phase Two and Assessment Five at follow-up, after a period of no intervention.

At each of these assessment times the set of data outlined below were collected: picture naming, recording a conversation and completing a number of language control tasks. The picture naming test monitored performance in the set of 200 pictures, subsets of which constituted the treatment and no treatment sets in therapy in both phases (see details below). This enabled comparisons to be made between performance prior to and after therapy. In order to investigate any carryover to natural speech situations a conversation was conducted between the person with aphasia and a chosen conversation partner. Finally a set of language control tasks was used which monitored performance in language functions not thought to be directly addressed by the therapy.

The first stage involved a period of language assessment and an experiment investigating facilitation of picture naming of target words. The assessment stage began with informed consent being obtained, and included tests of auditory comprehension, expressive language, reading and writing, and cognitive skills.

This was followed by the first phase of therapy (Phase One) which focused on improving word-finding in a picture naming task. Participants were seen once per week for a total of eight weeks, each session lasting roughly one to two hours. After this participants were invited to enter the second phase of therapy (Phase Two) which aimed at enabling the person with aphasia to use treated words in tasks approximating more closely to real-life conversation. After this phase of therapy participants were not seen for two months then were reassessed to determine the maintenance of any therapy effect.

5.2 PICTURE NAMING ASSESSMENT AND OTHER OUTCOME MEASURES

5.2.1 Set details

The pictures used in this study were compiled for the study from other sets including Nickels' (1992) European Naming Test (unpublished). The overall set contained 200 black and white line drawings of pictureable items. Within the total set matched subsets were incorporated, to allow the analysis of the impact of key variables on picture naming. The variables incorporated into the matched subsets were: imageability, animacy, operativity, word familiarity, word frequency and familiarity combined, age of acquisition, and number of syllables. Many of the values were taken from the MRC Psycholinguistic Database (Coltheart, 1981). Frequency values were taken from the

Celex database (CELEX, Centre for Lexical Information, 1993). See Howard Best Bruce and Gatehouse (1997) for details of rating operativity.

5.2.2 Naming test procedure

Each picture was presented individually to the person with aphasia, and they were asked to find the best word to describe the picture. Responses were tape-recorded and transcribed in situ. When the person with aphasia produced a string of responses, the last response within the first five seconds of the picture being presented was scored. Timings were carried out manually with a stopwatch.

TABLE 5.1 Design of the study

Assessment One: Baseline 1 (naming, conversation, language control tasks)
Eight week period of assessment of language functions Facilitation experiment (Best et al, 2002)
Assessment Two: Baseline 2 (naming, conversation, language control tasks)
Eight week period of therapy: Phase One Lexical therapy (Hickin et al, 2002)
Assessment Three: Post Phase One assessment (naming, conversation, language control tasks)
Eight week period of therapy: Phase Two (Herbert et al, 2003) Combines lexical and communicative approach
Assessment Four: Post Phase Two assessment (naming, conversation, language control tasks)
Eight week period of no intervention
Assessment Five: Follow-up assessment (naming, conversation, language control tasks)

5.2.3 Response categorisation

The transcribed responses were then analysed and assigned to categories as follows:

- correct responses included the target and other acceptable responses, such as a close synonym of the target, and minor phonetic errors;

- visual errors included misperceptions of the picture stimulus, and naming of a part of the picture instead of the whole;
- semantic errors included superordinate, co-ordinate, subordinate, and associative errors, circumlocutions, and words with a semantic and other relationship to the target, notably a visual relationship or a phonological relationship;
- phonological errors were words or non-words sharing at least 50% of the target's phonemes;
- unrelated word and non-word responses, and morphological errors;
- rejections where the person clearly rejects their own response as incorrect;
- no response where there was a complete failure to respond or the person made comments about their lack of response.

The criteria used here were drawn up by Best, Bruce, Gatehouse and Howard (Wendy Best, personal communication).

5.2.4 Written naming

A subset of 40 of the items used for the picture naming test were used to assess written naming. This subset was controlled for word length, frequency and imageability. Items were presented in the same way as for picture naming. The person's best attempt was scored with no time limit.

5.2.5 Conversation

In order to measure any effect of the two therapies on conversation, participants were asked to tape record a normal conversation with a family member or a friend. The development of a reliable means of measuring change in conversation, and the methods used here, are covered in depth in Chapter Six. In order to participate in the study the person with aphasia had to identify a conversation partner who would be available for each assessment point. They were then asked to tape-record an everyday conversation with this person, of 10 to 15 minutes in length at Assessments One to Five. These tape recordings were then analysed by the author of this thesis and a colleague. Key variables pertaining to noun production in conversation were identified, and these served as outcome measures for therapy.

5.2.6 Language control tasks

A set of language control tasks were identified and were repeated at each assessment point. These were language functions which it was predicted would not change as a result of a noun production therapy. They were: written sentence comprehension (CAT: see section 5.3.6); picture pointing span (see section 5.3.3), reading aloud words (see section 5.3.5) and reading aloud non-words (see section 5.3.5). Where a participant was at floor on a particular test that test was administered in the auditory rather than written modality.

5.3 LANGUAGE ASSESSMENTS

5.3.1 Tests of semantic processing

Semantic processing was tested by means of three assessments: Pyramids and Palm Trees three picture version (Howard and Patterson, 1992); spoken word to picture matching and written word to picture matching (CAT Comprehensive Aphasia Test: Swinburn, Howard, and Porter, in press). This allowed semantic processing to be investigated via three independent access routines.

5.3.1.1 Pyramids and Palm Trees

This test consists of 52 test items and includes norms for an adult population. Normal adults' mean score was 98 - 99% correct, with a minimum score of 49/52. The authors report that a score of 90% or more indicates no significant deficit. Each test item consists of a stimulus (e.g. glasses), the test target (e.g. eye), and a distractor which is a semantic co-ordinate of the target (e.g. ear). The examinee is required to point to the target. Success in the task indicates intact semantic knowledge about the test item's relationship to the semantic associate. In all cases the semantic relationship is associative but in some cases a logical inference must also be made in order to perform that task correctly (e.g. stimulus bat, target owl, distractor woodpecker). A further issue concerns the ethno-centricity of the task, relying as it does on English/European semantic knowledge e.g. concerning pigs' predilection for acorns. It can be performed in several ways: using three pictures, thus not testing language processes directly; using three written or spoken words, or using a combination of verbal and pictorial input. In

the study reported here the three picture version was used. This allows semantic processing to be investigated without recourse to any verbal processing and thus allows disparities between the three input routines to be identified.

5.3.1.2 Word to picture matching

Word to picture matching was investigated using the relevant subsets of the CAT. Here a target word must be matched to one picture from an array of four. The three distractors include a semantic distractor, a phonological distractor, and an unrelated distractor. The test exists in spoken and written forms and both were performed here. Analysis of the pattern of errors allows conclusions to be drawn regarding input processes: primarily semantic errors would indicate a semantic impairment, primarily phonological errors would indicate damage to auditory or visual input processes accessing semantics.

5.3.2 Auditory input

The Auditory Discrimination subtest from the Action for Dysphasic Adults Comprehension Battery (Franklin, Turner and Ellis, 1992) was used to test pre-lexical phoneme discrimination at the level of auditory analysis. This was felt to be important as the facilitation and therapy phases relied on accurate perception of sub-lexical auditory input cues in the form of initial phonemes.

The test consists of 40 pairs of CVC structures, half of which are identical (e.g. /kɛd/ /kɛd/) and half of which differ by one phoneme from each other (e.g. /kɛb/ and /kɛd/). Participants are presented with the two non-word stimuli and are asked to determine whether they are the same or different. Presentation is by tape recorder, and the two stimuli are presented by in turn one male and one female voice randomly ordered across stimuli. In the set of non-identical pairs of stimuli the difference between the two is located in either the first phoneme, the last phoneme, or the vowel. The consonants vary in terms of the number of features by which they differ, one or two features.

5.3.3 Auditory short term memory

Four tasks were used to assess the function of auditory short term memory. It was important to test this as the choice element introduced into the facilitation and therapy method involved the person with aphasia holding a number of stimuli (the correct initial phoneme plus up to three distractors) in short term memory. An interesting analysis might involve the relationship between the ability to respond to the choice of cues, and the function of auditory short term memory.

The tasks involved exposure to strings of stimuli, and either the repetition of these (digit span, letter span, and phoneme span), or pointing to pictures (picture span). In the first three tasks the person heard the stimuli (e.g. "two, five, three") and repeated these. In the latter method the person heard object names (e.g. "star, house, pen, tree") and pointed to the pictures in the correct order. Ten trials were performed in each task. If the person was successful at one level in terms of correct items produced in the correct order, the next string was increased by one. If unsuccessful, the next string was decreased by one. This gave ten attempted trials. The first trial was discounted, and a projected eleventh trial included, and the mean of these ten trials was computed. This gave a score for that task.

5.3.4 Internal phonology

In order to identify the level of breakdown in picture naming, it is desirable to include tests of internal phonology. These tests are assumed to tap knowledge in the phonological output lexicon. A person may be able to access knowledge about a word such as its first phoneme, or its rime²².

Two tasks were used to investigate internal phonology: an initial phoneme judgement task (Best, unpublished), and a homophone judgement task (Gatehouse and Best, unpublished). The initial phoneme task requires the person to consider three pictures, and determine which two begin with the same sound. Letter-sound regularities are varied in the task such that there is not always a direct correspondence between first letter and first sound, with items such as 'knee' being included. The homophone task

²² This is similar to the knowledge a non brain damaged person may have about a word when in a tip of the tongue state.

requires the person to determine which two of a choice of three pictures sound the same (e.g. flour, flower, bed), where two of the pictures represent the two meanings of the pair of homophones. Participants are asked to perform both tasks silently.

5.3.5 Tests of output phonology

Reading aloud and repetition of words and non-words were used to investigate output phonology. The same sets of words and non-words were used for both tasks.

Participants were shown a written stimulus or heard a spoken stimulus. In the word set there were 182 items. This set was controlled for frequency, imageability, and length. The non-word set consisted of 26 items, controlled for length (both word and non-word sets were compiled by David Howard, personal communication). There was no relationship between the set of items used here and those used in the picture naming test.

5.3.6 Sentence comprehension

The sentence comprehension subtests from the CAT (Swinburn et al, in press) were used to investigate syntactic processing. The test exists in spoken and written forms. There are 16 stimuli in each sub-test and they probe a range of syntactic structures. The stimuli include simple subject verb and subject-verb-object structures, passive structures and embedded clauses. The latter are likely to be problematical for people with aphasia.

5.3.7 Cognitive skills

A set of cognitive tasks were administered in order to identify any participants with significant cognitive impairments likely to impact upon their ability to complete the assessment and therapy tasks included in the study.

5.3.7.1 Non verbal problem solving: Raven's Progressive Matrices (Raven, 1967)

This task involves visual problem solving. Participants are presented with a coloured visual design, from which one section has been removed, rather like a hole cut in a piece of material. Below this there are six stimuli, the target and five distractors, from which the participant has to select the correct missing part. This task requires fine visual skills

in terms of line, shape, and colour discrimination, and visual problem solving in determining which part would fill the empty space appropriately.

5.3.7.2 Visual processing: Object decision

The object decision task (Birmingham Object Recognition Battery, Riddoch and Humphreys, 1992) offers a closer view of visual processing of pictorial material. This was administered in order to identify any participant with significant visual processing deficit likely to affect their ability to derive meaning from pictorial information. The stimuli consist of 32 items, half of which depict real objects and animals and half of which depict nonsense objects and animals. The participant has to determine through a yes/no decision which are real and which are nonsense.

5.4 FACILITATION OF SPOKEN WORD PRODUCTION

5.4.1 Design and procedure

Two forms of facilitation were investigated as these had a direct relationship with the therapy to be carried out in Phase One: spoken first consonant and vowel and written first consonant and vowel were used. The test items were a set of 164 black and white line drawings representing words with a CVC segmental structure. None of these words appeared in the 200 word picture naming test. All had at least 90% name agreement from normal controls and had single syllable CVC names. The mean log frequency was 1.45 (s.d. 0.56, range 0.18 to 2.78; combined spoken and written frequency from Celex database, 1993).

Participants were presented with a set of pictures to name. If they were unable to name a picture within five seconds it was entered in one of three conditions: extra time, which acted as the control condition, single cue or choice of two cues. This procedure is shown in Figure 5.1. Items were assigned in sets of four in rotation to the three conditions until there were a total of 12 items in each condition. Thus there were 36 items overall.

Thus, having failed to name an item in the first five seconds, participants were either given a further five seconds to name the picture (extra time) or they were given a cue, or choice of two cues. They were required to attempt to name the picture in the five seconds immediately after presentation of the cue.

The two cue types were presented in two separate sessions. The two sessions containing facilitation took place at least one week apart. The cue types were:

- spoken CV in which the first consonant and vowel of the target were spoken (e.g. /bɔ/).
- written CV in which the first consonant and vowel of the target were presented in upper case, 18 point, written form (e.g. BA).

In the choice condition the target cue was paired with the comparable information for an unrelated word. For example, in the spoken CV condition if a picture of a cat was not named within 5 seconds and was allocated to the choice condition it might be cued with 'begins with /bɔ/ or /kæ/'. The order of the correct and distractor cues was randomised. The distractor cues had no phonological overlap with the target cues, nor did they constitute the first sounds or letters of a semantically related word.

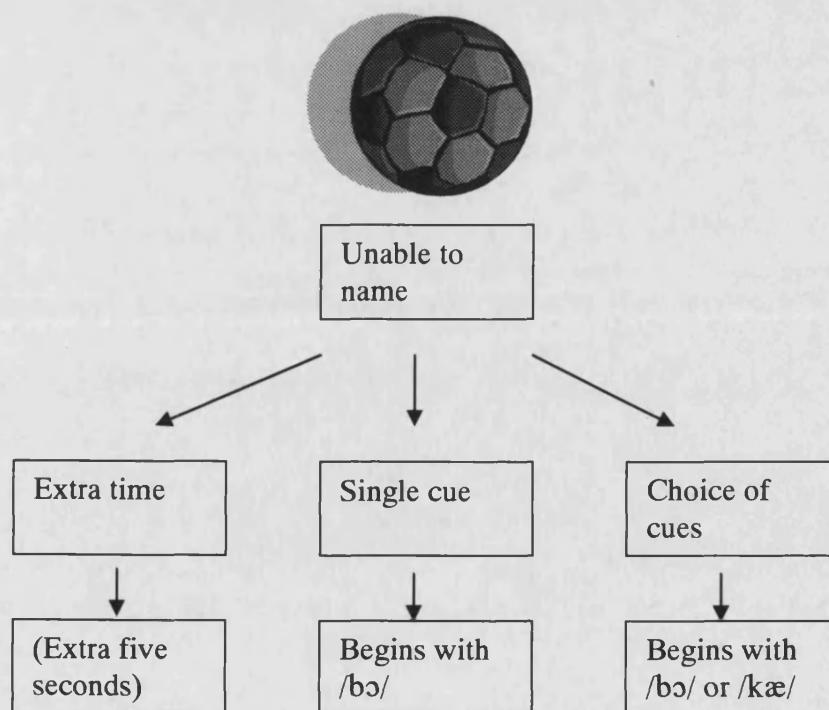


FIGURE 5.1 Design and procedure for naming facilitation (CV spoken condition)

5.5 THERAPY PHASE ONE

5.5.1 Design

Items undergoing therapy in Phases One and Two were taken from the total set of 200 pictures, which participants named at assessments one to five throughout the study (see Table 5.1). From the results of the naming assessments one and two (prior to Phase One), a subset of 100 words were selected for therapy for each participant. Thus all participants were tested on the same items at assessments one to five, and all therapy items were taken from the same set of 200 words, but the items selected for therapy differed across individuals. The remaining 100 items acted as untreated controls. The 100 treated items were split into two sets of 50. One set received phonological therapy, and the second received orthographic therapy.

Selection of the treated and untreated items for a given individual was performed by splitting the 200 items into four sets: those named successfully at assessments one and

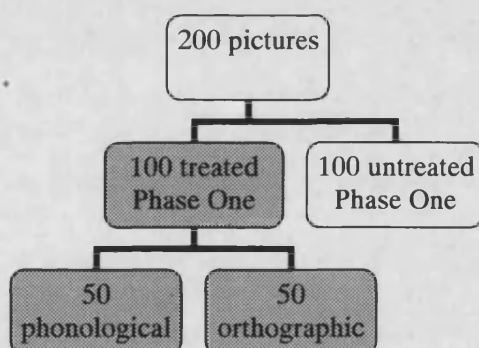


FIGURE 5.2 Phase One Therapy:
Allocation of items to therapy and control sets

two, those named only at assessment one, those named only at assessment two, and those not named correctly at either assessment one or assessment two. Each of the four sets was then split randomly into two, half going into the therapy set and half going in to the untreated set. Thus treated and untreated sets were matched for performance at baseline. The treated set was then further split into two sets, phonological and orthographic, using the same procedure outlined for the treated and untreated sets

above. Thus the phonological and orthographic sets were matched for baseline performance. To summarise, any given item was either treated or untreated in Phase One, and if treated, received either phonological or orthographic cueing. Allocation of items to sets is shown diagrammatically in Figure 5.2.

5.5.2 Therapy Phase One Procedure

The treatment administered in Phase One falls into the category of phonological approaches in that it sought to highlight phonological or orthographic information for the target word. In this respect it differs from the majority of therapy studies reported so far in the literature, which highlight the meaning of the word, or the meaning and the form, but less commonly the form in isolation.

The procedure was as follows: participants were shown a picture to name, and if they could not name it within the first five seconds of seeing the picture, or they made an error in attempting to name it within that time, they were then exposed to the first sound of the word in the phonological condition or the first letter of the word in the orthographic condition, with one or more distractors. Targets and distractors were matched for segmental structure, and as far as was possible for orthographic form as well, but they differed in terms of the first letter and sound, and first vowel. Distractors were not semantically related to the target and as far as was possible distractors were controlled to avoid inadvertent cueing of a semantic relative of the target (e.g for target “cat” a distractor beginning with “d” would not be included in case “dog” was inadvertently cued). Target cues and distractor cues appeared in random order. For the first two sessions one distractor was present. These increased to two for sessions three and four, and then to three for the final four sessions. The procedure for the Phase One therapy is shown in Figure 5.3.

The procedure for the orthographic and phonological conditions was identical, except that in the phonological condition the person heard the spoken cues and in the orthographic condition they saw the written cues. Each of the 100 therapy items was treated once per week in a single session lasting roughly one hour. The sessions totalled eight in number. Although for the majority of participants the therapy procedure was consistently adhered to, there was room for individual adjustment, for example for a

person who had difficulty with more than two distractors. All three participants included in the study reported here were however able to cope with the full set of distractors.

5.6 THERAPY PHASE TWO

5.6.1 Design

A further 100 set of items were selected for therapy in Phase Two. From the set of 200 items, 50 which had received therapy in Phase One, and 50 which were untreated controls in Phase One were assigned to the new therapy set for Phase Two. This left a further 100, half of which had received therapy in Phase One, half of which had not. These items acted as untreated controls for therapy Phase Two. Thus at the conclusion of Phase Two, 50 items had been treated in both Phases, 50 had been treated in Phase One only, 50 had been treated in Phase Two only, and 50 had received no therapy in either Phase One or Phase Two.

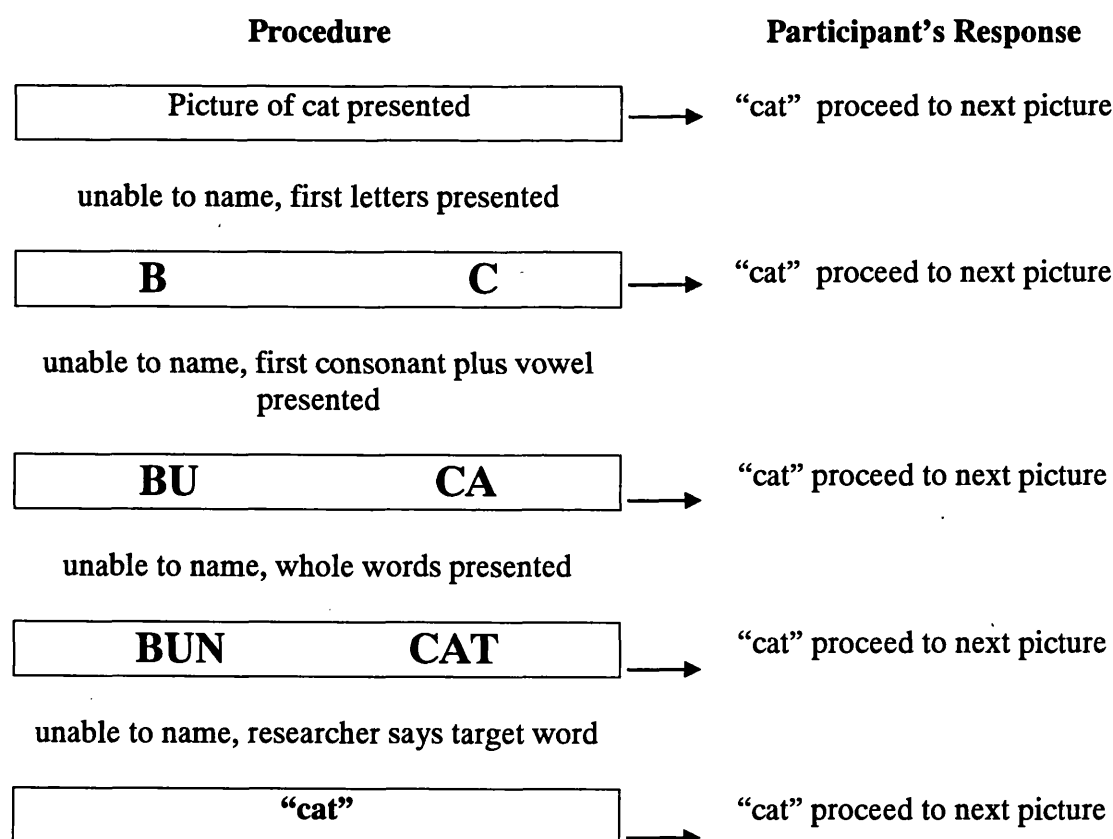


FIGURE 5.3 Procedure for Phase One Therapy Orthographic Condition

For Phase Two selection of functionally relevant items occurred, so items were assigned to the treated and untreated sets according to how relevant the word was to the individual's communication situation. This differed from Phase One, where items were selected purely according to their performance in the baseline assessments. Sets were nevertheless matched for performance across the three previous assessments. The allocation of items to sets in both Phases One and Two is shown diagrammatically in Figure 5.4.

5.6.2 Phase Two Therapy Procedure

The resulting 100 picture items were sorted into conversational categories, such as shopping, family, household items, and holidays, and made into a file which the participant kept in their home throughout the therapy phase. Each item consisted of a line drawing with the written cues for the target word provided below. The written cues included the target word plus two semantically unrelated words which were matched for segmental and orthographic structure as closely as possible (e.g. Target: milk. Distractors: gold and desk). On the first line beneath the picture the first letter of the target and distractors was provided. On the next line the first letter and the first vowel were provided, and on the third line the total word for the target and the distractors. Thus participants could self-cue using the written cues by gradually revealing more and more of the target as needed. In the therapy sessions phonological cues were also offered in addition to the orthographic information if the participant required.

The emphasis in the therapy was on the use of the target words in everyday speech. To this end tasks used in the therapy sessions approximated to real situations as far as possible. The therapy proceeded in a hierarchical fashion, moving gradually away from picture naming, towards natural use of words in conversational speech. In the early sessions all 100 words were treated in each session; in the later sessions subsets were treated depending on the conversation topic in question.

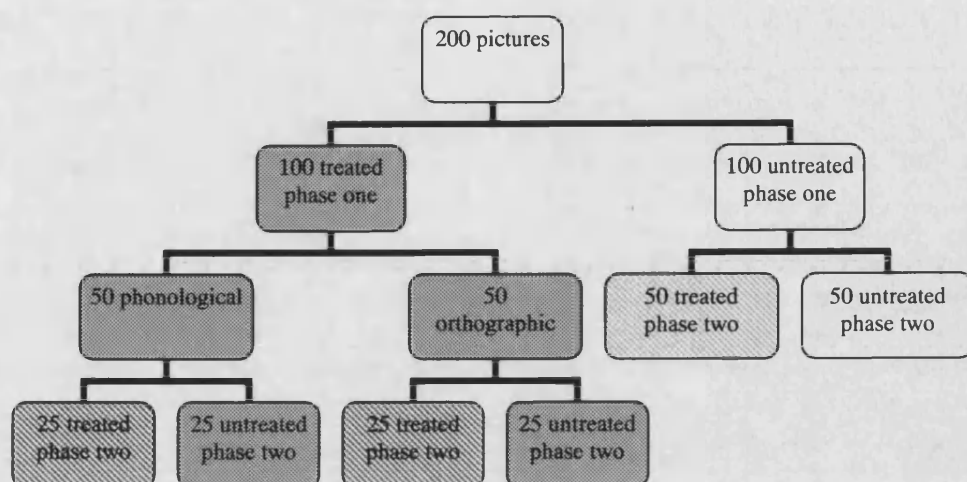


FIGURE 5.4 Phase Two Therapy - Allocation of items to therapy and control sets

In the first two sessions two tasks were used with 50 items allocated to each task. For 50 items a spoken definition was presented for the person to name; if they were unable to do so the picture and orthographic cues were available to help the person but there was no pressure from the researcher for the participant to achieve the target word. For the second set of 50 items picture naming was conducted in a pseudo-realistic speech situation similar to that described in PACE (Davis and Wilcox, 1981) and based on Pulvermuller and Roth's (1991) request game. Participants were asked to communicate to the researcher the meaning of one of the set of picture items they held before them. The researcher held the same set, and would identify from the participant's communication attempts which item they were referring to. Thus in the first two sessions each item was exposed to both interventions (naming to spoken definition, and the Request Game) once.

In the third and fourth sessions participants made spoken lists of treated items according to the task in question (e.g. shopping list, household inventories, aide-memoires, holiday requirements, Christmas presents). Lists were selected according to the participants' interests and the contents of the therapy file. In sessions four to eight participants were free to select topic areas to talk about and the sessions involved conversations with the researcher. These involved reminiscing, telling anecdotes, making plans for the future, and were based on the conversation topics identified in the therapy file. In all sessions participants had access to the pictures, and to the written

cues, should they need them, and were encouraged to refer to the file when a word-finding difficulty arose. Prior to each conversation the participant and researcher looked through the set of possibly relevant items and the participant was free to attempt to name them if they wished. It is worth noting that in sessions three to eight the person was not exposed systematically to all of the therapy items (this differs from the Phase One therapy where each item was seen in each session). A record was kept throughout the therapy sessions of words produced by the participant

5.7 SUMMARY

In this chapter the assessment and therapy stages of the study have been outlined. The overall design of the study and the method for assigning items to treated and untreated sets have been described. In Chapter Four details of design issues relevant to therapy studies were outlined (section 4.2.3.1). In the study described here a number of design factors were implemented to ensure that the study met the necessary criteria for establishing the efficacy of a method. All of the participants were over one year post onset and therefore were not likely to be experiencing spontaneous recovery; to control for a generalised therapy effect a set of language control tasks were implemented; and to measure generalisation beyond therapy items matched sets of treated and untreated items were used.

Of the outcome measures used in the study one looks at picture naming, and one looks at connected speech in conversation. As has been discussed in Chapter Four there are difficulties in using conversation as an outcome measure unless the reliability of the various factors being quantified has been established. In the next chapter details of the development of a means to reliably quantify conversational behaviour is described.

CHAPTER SIX: DEVELOPING A MEASURE OF WORD FINDING IN CONVERSATION

6.0 INTRODUCTION

For many people with aphasia word-finding difficulties form a significant part of their language disability. Lesser and Algar (1995) and Perkins, Crisp and Walshaw (1999) found that word retrieval problems were likely to lead to conversational breakdown and subsequent collaborative repair for the majority of the people with aphasia they studied. For the person with aphasia word-finding difficulties can thus translate into a real and significant handicap in social interaction, often making conversation with friends and family laborious or unsuccessful with speakers failing to achieve mutual understanding. In some cases conversation is avoided by one or both parties, resulting in suspended or curtailed relationships and the social isolation of the person with aphasia.²³ For these reasons, rehabilitation of word-finding deficits in aphasia must result in more than improvements in picture naming, encouraging though these may be, and must be able to demonstrate real effects in terms of the quality of the exchange of meaning in conversation. It follows that the goal of aphasia therapy, regardless of the type of language deficit, will often be to improve conversational abilities to an optimum level.

The issue of establishing carryover of the effects of impairment-based therapy to everyday speech remains one of the challenges facing therapists and researchers working with people with aphasia. As was outlined in Chapter Three (section 3.4), few studies have attempted to measure this, and those that have done so have limited themselves to investigating connected speech in tasks such as story telling. Such tasks are vulnerable to the same criticism leveled at picture naming itself: namely, that they do not measure performance in everyday conversation, and thus lack ecological validity.

In order to address this issue directly a means of quantifying word finding behaviours in conversation was developed in this study. The measure addressed both successful word-finding and word-finding breakdown, looking at the effect of these on the nature of and flow of the conversations. In this chapter the background to the measure is described, followed by analyses of the reliability of the measure.

²³ One possible effect of aphasia is that of social exclusion. See e.g. Parr Duchan and Pound (2003).

6.1 MEASURING CONVERSATION

One overwhelming difficulty in this line of investigation concerns the relatively unconstrained nature of conversation, with two conversations between any two speakers showing variation across a number of parameters. Boles and Bombard (1998) demonstrated an extreme example of this variation in their study of repair in conversation. Such wide variation may exist in all aspects of conversation, given that speakers are almost entirely at liberty to behave as they wish. As a result the use of conversational data as an outcome measure is problematical. Both qualitative and quantitative differences found between any two conversations within a given pair of speakers may be part of this inherent variation. Additionally many aspects of conversational data are open to interpretation by the analyst. It is feasible however, that certain aspects of speech production within conversation, such as the number of neologisms an individual produces, will be stable across independent conversations.

A number of studies have attempted to establish the reliability of conversational data. In most cases a percentage of the data was analysed by two raters. Boles and Bombard (1998) looked at transcription agreement: a second judge transcribed and coded 20% of all conversational interaction in nine conversations from a total of 44. They found inter-rater reliability of 0.96 for transcription, 0.95 for utterance boundaries, and 0.86 for repair and intra rater reliability of 0.99, 0.99, and 0.95 for the same variables. Watson, Chenery and Carter (1999) looked at inter rater and intra rater reliability for trouble indicating behaviours, repair trajectory and repair type and found over 0.90 agreement for all variables for 30% of the data. Oelschlaeger (1999) analysed raters' agreement in identifying repair sequences and found intra rater reliability of 1.00 and inter rater reliability of 0.88 for a subset of the data. In most of these studies inter-rater (two independent raters considering the same piece of data) and intra-rater agreement (the same rater considering one piece of data at two separate times) was sought.

It is important that reliability studies assess inter-rater and intra-rater agreement but this does not address the variation across conversations. The latter necessitates test-retest stability measures. It was therefore considered essential in developing the measure

described here that inter-rater and intra-rater reliability, and test-retest stability were sought for all conversational factors measured. In addition measures of agreement which consider only the proportion of items coded in the same way fail to take account of chance. In the analysis described here the proportion of agreement is considered but in addition the Kappa statistic (Cohen, 1960) is computed where possible as this incorporates an adjustment for chance coding of a behaviour into a category.

6.2 OVERVIEW OF THE RELIABILITY STUDY

The measure designed for this research aimed to quantify successful word retrieval and failures in word retrieval and the effects of these on conversations between the person with aphasia and their partner. In the initial phase of the study inter-rater and intra-rater reliability of selected aspects of conversation data was investigated with five conversation dyads, and as a result of this initial analysis aspects which showed poor reliability were eliminated. In the second phase of the study test-retest stability of the resulting measure was investigated with the first five dyads and a further five conversation dyads. Variables which showed good test-retest stability were then included in the final measure. The final measure is the result of these two phases of analysis.

The data and analysis therefore consist of the following:

Stage one:

- analysis of inter-rater and intra-rater reliability;
- draft one of the measure was used;
- five conversation dyads took part;
- conversations are from assessment one;
- inter-rater reliability analysis: two raters analysed the data (RH and another);
- intra-rater reliability analysis: one rater (RH) analysed the data on two separate occasions.

This analysis resulted in draft two.

Stage two:

- analysis of test-retest stability;
- draft two of the measure was used which included variables showing satisfactory levels of inter-rater and intra-rater reliability;

- ten conversation dyads took part (five from stage one plus five others);
- conversations were from assessment one and assessment two;
- the analyses were performed by one rater (RH).

This analysis resulted in draft three, the final version.

6.3 DEVELOPMENT OF THE MEASURE

6.3.1 Background and sources

In recent years methodology developed in the field of conversation analysis (CA) has been applied to the study of the interaction between people with aphasia and their conversation partners. In particular the study of how conversations get into difficulties and how the two participants engage in sorting out the problems has provided fruitful data. In devising the measure attention was paid to the research carried out in the field of CA, and the measure drew on the findings of such studies. It is important however that the measure quantify conversational behaviours and not merely describe these. Proponents of CA caution against quantification however: Schegloff (1988) and Lesser and Milroy (1993) propose that quantification ignores the function of turns in the sequence of talk. In contrast, Heritage (1999) wrote “I want to consider the likelihood that CA will become more quantitative during the next period of its development” (Heritage, 1999: 70) indicating that this is a route now under consideration by certain CA proponents.

The measure reported here owes a lot to the work done by Crockford and Lesser (1994)²⁴. In their analysis of two conversations between people with aphasia and their conversation partners on separate occasions, they quantified the following: the number of editing elements (after Schlenk, Huber and Willmes, 1987); the instances of collaborative repair; and the weight of conversational loading as measured by the conversational partner’s ratio of minimal responses and initiations of information.

Although initial attempts were made to count ‘initiations’ in the measure described in this thesis, this was abandoned as raters reported an inability to decide what these were (see section 6.4.4. below). In the measure described here editing elements were

²⁴ Although in Crockford and Lesser’s study inter-rater reliability and test-retest reliability were not analysed

included, along with the types of turns employed by the speakers and the amount and nature of the breakdown and repair. In addition, successful lexical retrieval and topic initiation were included in the measure.

6.3.2 Type of sample

As the aim was to measure what is happening in real life conversation as closely as possible it was felt appropriate to avoid the institutional interactions that can occur between professionals and people with aphasia. Lay-professional conversations are known to differ in significant ways from those not involving a professional. For example Perkins (1995) found that in conversations with the researcher the aphasic speaker produced more major turns than in conversations with their partner. Such conversations may not therefore be truly representative of the person's everyday conversation. The measure therefore samples conversations tape-recorded between the person with aphasia and their chosen conversation partner, in most cases a friend or spouse. To use Holland's term this data has strong 'ecological validity' (Holland, 1991: 199). For the measure described here the same conversation partner participated throughout, thus eliminating a further possible source of variation. The person with aphasia had known the conversation partner for a number of years, and in most cases since prior to the stroke. The topic of conversation was not pre-determined: participants were merely asked to carry out a normal conversation.

6.3.3 Recording format

In any observational research the introduction of the observer necessarily alters the nature of the behaviour. This is known as 'the observer's paradox'. In order to minimize this effect the observer must be as unobtrusive as possible. Many researchers into CA have used video-recordings of conversations to identify as closely as possible the nature of conversational interaction²⁵. This has clear advantages over audio- tape recordings in that non-verbal communication can be identified and measured, but the method is intrusive.

²⁵ Others who have used tape recordings include Perkins (1995), Oelschlaeger (1998), and Crockford and Lesser 1994.

Tape recordings were selected for the measure described here for three reasons. First the method is less intrusive, and it is easier for participants in the research to operate the technology. Second, the focus is on word retrieval in conversation, which by its very nature is audible and therefore likely to be captured in the recordings²⁶. Third, for the measure to be useful clinically it needs to be straightforward for clinicians to use.

6.3.4 Length of sample

As stated above, one of the main difficulties associated with measuring conversational data is that there is inherent variation across conversations between any two people. One example of this is where one speaker dominates one conversation, but not a second. Conclusions cannot be drawn easily from such data. It might be thought that the longer the conversations under consideration the more equitable the two conversations will be, but this raises the obvious practical issue of how long. The length of samples in published research varies. Perkins (1995) looked at 12-minute samples, whereas Silvest (1991) looked at five-minute samples. It was important however that the measure be feasible to use in a clinical setting where time is more limited. Following Silvest (1991) 10 -15 minute recordings were obtained. Kennedy et al (1994) describe the beginnings and ends of conversations as ritualized and therefore possibly unrepresentative, thus these were avoided and the middle five minutes formed the data sample.

6.4 THE CONVERSATION MEASURE: DRAFT ONE

The focus is on word retrieval, success and failure, and its effect on conversational interaction, notably on the degree to which word retrieval difficulties may cause breakdown in conversation and engender subsequent repair. As a result the measure concentrates heavily on these aspects of output. In addition, topic initiation was included in the first version of the measure. Draft one of the measure is included in full in Appendix Two.

²⁶ It is however the case that certain aspects of the conversation were lost to the analysis. In particular gestures and pointing demonstrating the meaning of the word being sought, and other non-verbal indications crucial to the interaction, were not available to the analysis.

6.4.1 Number of topics and topic initiator

The presence of a language disorder, in particular a word retrieval deficit may result in a speaker being unable to introduce topics of their choice. It was therefore felt important to try to measure the instances of new topic introduction and the distribution of new topic introduction across the two speakers. Perkins et al (1999) report having to abandon attempts to measure topics as they were unable to generate meaningful definitions (Lisa Perkins, personal communication).

6.4.2 Number of speech units

The total number of speech units for both speakers is counted. This allows a comparison between the person with aphasia and their conversation partner, thus giving a crude measure of the relative contribution of the two speakers to the conversation. In addition this measure serves as a denominator to allow proportions of occurrences of other factors to be computed, for example the proportion of speech units which are content words. Thus it is possible to distinguish between the amount of spoken output and the contribution of that output to the conversation. This is particularly important as people with aphasia vary greatly in the amount of speech they produce, differing along the dimension of fluency²⁷.

6.4.3 Number of turns

Thus was computed for both speakers, again allowing a comparison across the two speakers to be made. In transcription of the tape-recording turns were determined according to Silvast's (1991) criteria, which states that a new turn is indicated where there is a change of speaker, or a pause of 1.5 seconds or more after the completion of a semantically coherent contribution²⁸. Delays in the person with aphasia's speech initiation, while they search for a word or attempt to construct a sentence, may lead to their losing the floor more easily than a non-aphasic speaker (Lesser and Milroy 1993).

²⁷ It is possible for example that non-fluent speakers, who produce few speech units, may contribute relatively efficiently to the conversation, whereas fluent speakers may produce large numbers of speech units but demonstrate more breakdowns in shared meaning in their conversations.

²⁸ Within the people with aphasia's turns pauses of more than one point five seconds may occur due to word retrieval problems for example. The subsequent speech was not treated as a new turn if the preceding speech indicated the pause was due to a search for a word or a grammatical construction, and the other speaker did not start a turn at this juncture.

6.4.4 Types of turns

One of the problems in interpreting data from studies of conversation lies in clarifying what is meant by the terms major or substantive turn, initiation, and minimal turn. In many reports no definition is provided. Where definitions are provided these may differ across research groups. It was therefore important to clarify exactly what was included and excluded from the turn categories.

Two types of turns were identified. The aim was to identify the speaker's turns which contribute meaningful information to the conversation. Such turns have been variously described as substantive or major turns. This is one measure of the amount of conversational load taken by each speaker. Crockford and Lesser (1994) include a category of turns which they call initiations. Initial attempts to include this in the measure were abandoned as it rapidly became clear that raters were extremely unconfident about judging turns in this way. In attempting to measure turns it was essential to define exactly what was meant by a substantive turn: this was eventually defined as a turn containing at least one content word²⁹.

One phenomenon identified in turn-taking in conversation is that of the perverse passive (Jefferson, 1984) in which one speaker produces minimal utterances (such as *mmm*, or *yeah*) in response to conversational gambits by their partner, thereby signaling an inability or unwillingness to take the conversational floor. By such turns the speaker does not move the conversation on but merely signals acquiescence in its being continued by the other speaker. Perkins (1995) found that the time pressures of conversation may force the person with aphasia to contribute mainly minimal responses. The non-aphasic conversation partner may be intolerant to delays and self-repairs on the part of the person with aphasia, or they may take on the burden of the conversation in a mutually agreed fashion, in order to save face and to achieve some form of interaction. It is possible also that the more language-impaired a person is the more they have to rely on these turns to maintain their part in a conversation. It was therefore important to capture the occurrence of such turns.

²⁹ Where a content word is defined as: a noun, a verb, an adjective, an adverb ending in -ly, or a numeral (Bird and Franklin, 1996).

Such turns are also known as discourse markers (Lesser and Milroy 1993), or, within the field of discourse analysis, back-channelling (Ulatowska et al, 1992)³⁰. Comrie, Mackenzie and McCall (2001) defined minimal turns as tokens (*mhmm, yes/yeah, right, y'know, oh well, uhuh,, aye/och-aye*, used in isolation) and “single word responses to the researcher’s questions, comments or requests for clarification” (Comrie et al, 2001: 387).

In analysing these turns for the measure described here both the actual speech token produced, and their interactional function, were considered. Thus certain minimal speech items were not included in this category. For example, unlike Comrie et al (2001), a single word response of *yes, no*, or *don't know* in answer to a question was excluded from this category. This was because it was felt that these turns are often semantically sufficient and appropriate, and that they provide more in the interaction than merely handing the floor back to the other speaker, for example they can close a topic. In particular, when minimal responses were produced as part of a collaborative repair sequence, these were not included as minimal turns³¹. The resulting set of minimal contributions to the conversation were labeled minimal turns after Schegloff (1982) and Jefferson (1984) and were defined as minimal speech items such as *mmhm, yes/yeah, oh, oh well, oh dear, right, dunno*, or combinations thereof and exclusions were defined as *yes, no* when these occur in sequences of collaborative repair and in answers to questions.

6.4.5 Lexical selection

In order to measure the person’s overall access to content words the production of nouns, main verbs (i.e. excluding modal and auxiliary verbs), adjectives, adverbs, and numerals was counted (see Bird and Franklin, 1996, for a description of inclusion criteria). This gave a measure of success in lexical retrieval. In the category of lexical selection the number of speech errors produced was also estimated. These were labeled

³⁰ In Ulatowska et al’s (1992) study these turns were discounted from the analyses. In conversation they are however viewed as a crucial part of the aphasic person’s and their partner’s adaptation to conversation with aphasia.

³¹ In many cases of collaborative repair analysed here the sequence projected over many turns, with the conversational partner offering suggestions, which the person with aphasia either accepted or rejected. In these instances the interactional function of the minimal speech item amounted to more than merely handing back the floor and was, rather, a signal to the partner to continue guessing

‘trouble indicating behaviours’ by Schlenk, Huber and Willmes (1987), along with hesitation phenomena (pauses and filled pauses).

6.4.6 Trouble-indicating behaviours: Definitions

Schlenk et al (1987) cite the following in the category of trouble indicating behaviours: pauses, interjections (filled pauses), repetition of immediately preceding utterance, phonemic approximations, circumlocutions or semantic approximations, semantic paraphasias, phonological errors, neologisms, and comments. Pauses of two seconds or more within the person with aphasia’s turn were also included here. In the measure described here semantic paraphasias and circumlocutions were considered as one category. All trouble-indicating behaviours were analysed in the context of the conversation around them. This is particularly important in the case of semantic paraphasias, where the target may never be actually spoken but is assumed by the two speakers, or where the correct target may only become apparent in subsequent turns. In all cases each instance of the behaviour is counted as one event, thus multiple attempts at a target are each counted as separate paraphasias.

Semantic paraphasias were defined as a word which is semantically related to the stated or presumed target, and a circumlocution as a phrase or sentence indicating a stated or presumed target. Phonological paraphasias were defined as words or non-words which contain half or more of the phonemes of the target in the correct order, or where half or more of the phonemes of the error are present in the target in the correct order. Also included here were part word repetitions such as false starts. Neologisms were defined as non-words where a target was not apparent. Overuse of pronouns was defined as an instance of pronoun use where there was no clear referent. Comments were defined as overt comments on a word-finding difficulty such as “I can’t think of the name”. Repetition involved the repetition of an immediately preceding word or phrase within a turn.

In addition to overt manifestations of word-finding difficulties a person with aphasia may also show covert signs in increased hesitancy and pausing within conversation. A certain length of pausing is tolerated within normal conversation (Sacks et al, 1974 report this to be about one second). Pausing was defined here as any pause over two seconds in length attributable to the person with aphasia. This included therefore those

pauses which occurred within the person with aphasia's speech and those which were attributable to them but were not surrounded by their speech, for example where they may be the designated next speaker (when the previous speaker has asked them a direct question) but fail to take the turn within two seconds. In addition to silent pauses filled pauses such as *er* and *um* were counted separately.

6.4.7 Repair

6.4.7.1 Collaborative repair

Boles and Bombard (1998: 550) describe repair as “an attempt to modify one's own or the other person's utterance for the purpose of clarification”. As stated at the opening of this chapter one of the main contributors to aphasic conversational breakdown and subsequent repair is the occurrence of a word-finding difficulty. According to Schegloff, Jefferson and Sacks (1977), although breakdown in normal conversation does occur it is generally resolved within three turns. Moreover, there is a preference within normal conversation for speakers to repair breakdowns themselves within the current turn (termed self-initiated self-repair). This is preferred for a number of reasons one of which has to do with saving face amidst the socially delicate business of handling an apparent incompetence in speaking. People with aphasia also carry out self-initiated self-repair, which may involve partial attempts at a word, rephrasing, pauses and fillers (see e.g. Laakso, 1997). Laakso and Klippi (1999) found that the person with aphasia usually attempts self-repair initially, and if this fails they then go on to establish “a collaborative framework by clearly shifting the orientation to a co-participant” (Laakso and Klippi, 1999: 360). This then opens the floor to the conversation partner to participate in the repair.

Conversations between an aphasic person and their non-aphasic partner differ from non-aphasic dyads, in that often when a breakdown arises the person with aphasia is unable to solve the problem alone (e.g. in searching for a particular word). In such instances considerable participation from the non-aphasic partner occurs. This results in sequences of talk in which both parties seek a successful resolution to the difficulty, by a variety of means, often extending over long stretches of conversation. These sequences have been referred to as ‘collaborative repair’ by a number of authors. Oelschlaeger and Damico (1998) refer to repair in aphasia as ‘joint productions’ (i.e.

collaborative sequences which are not viewed as breakdown and repair but as a means of interacting in their own right), and Laakso and Klippi (1999) reported that in conversations where there are aphasic participants problems are often treated “sequentially in a collaborative fashion” (Laakso and Klippi, 1999: 345).

There is evidence that repair is more frequent in conversations involving people with aphasia. Ferguson (1994) found more instances of repair in conversations between a person with aphasia and a non-aphasic partner than in conversations between two people without aphasia, and more other-repair in the former. The relationship between the severity of the language impairment and the incidence of repair is complex. Oelschlaeger (1999) described collaborative repair in one couple and reported that repair “was not dependent on the unique skills or abilities of either speaker” (Oelschlaeger, 1999: 69). Repair in aphasia differs from that found in non-aphasic conversation in that long sequences of turns are often devoted to it, but also in the form that the sequence of turns takes. One form of repair witnessed in aphasic conversation involves the person with aphasia giving clues or hints as to the word being searched for and the job of the partner is then to guess what the solution is. Lubinski, Duchan & Weitzner-Lin (1980) term these ‘hint and guess’ sequences between the person with aphasia and their conversation partner.

The number of turns taken to repair the breakdown may also be of interest. One difficulty in estimating repair length relates to identifying the onset of the repair sequence. Wilkinson (1995) reported that breakdown can occur long before repair is initiated, for example a misunderstanding can be traced back several turns, but repair begins only when it is clear that there has been such a misunderstanding.

Milroy & Perkins (1992) compared the endings of repair sequences in aphasic conversations to endings of normal conversations where various moves are made in sequential fashion to close down the sequence (this may serve to establish an acceptance that the repair is completed satisfactorily). This is an unusual feature of aphasic repair not found in normal conversations (where repair is usually dealt with rapidly without calling attention to it).

6.4.7.2 Successful or unsuccessful repair

Even by devoting large tracts of conversation to repairing a breakdown speakers do not always reach mutual understanding. In many instances of collaborative repair in aphasia resolution of the problem is not reached. There are a number of factors which may contribute to this: the knowledge both parties have of the topic; the knowledge both parties have of each other, in particular the strategies developed to handle the aphasia; the ability of the person with aphasia to find alternative forms to communicate the intended meaning, either through spoken language or other methods; the overall severity of the person's aphasia. Perkins (1995) considered the amount of shared knowledge of the interlocutors to be important as this may enable a partner to maintain a conversation in which a researcher would have had to instigate repair. It may also make interpretation of conversations difficult - what is a breakdown to the outsider may be a completed conversation to the two speakers concerned. These difficulties notwithstanding, it was felt important to not only gauge how often repair takes place but also how successful are those attempts at repair.

6.4.7.3 Type of repair

Lesser and Algar (1995) described some of the main methods speakers use to initiate repair, in their study of the conversations of a small group of people with aphasia. These methods form the basis of the repair types used in the measure described here. In the first type the person with aphasia asks for help and the pair jointly reach the target. In the second type the conversation partner asks for and receives clarification. In the third type the conversation partner corrects a mispronounced utterance. These all result in successful repair. In the fourth type the conversation partner makes a guess at the target and this is not resolved. In the fifth type the conversation partner's request for clarification is not resolved. These latter two are instances of unsuccessful repair.

6.4.8 Summary of draft one

In this section the rationale for the inclusion of the selected variables in draft one of the measure has been outlined. This first draft is shown in Appendix Two. Difficulties in compiling this first draft have been indicated and key definitions for certain variables provided. In stage one of the reliability study the first draft of the measure was used to

analyse data from five participants with aphasia. Inter-rater and intra-rater reliability of the measure was tested. Details of this first stage are outlined below.

6.5 INTER AND INTRA –RATER RELIABILITY

6.5.1 Method

6.5.1.1 Participants

Five people with aphasia took part in the first stage. Three of the five (HM, PH and SC) went on to take part in the subsequent therapy phases of the study (reported in Hickin et al, 2002 and Herbert et al, 2003), and two of those three (PH and SC) are reported in depth in this report (see Chapters Seven, Eight, and Ten). Details of the five participants and their conversation partners are given in Table 6.1.

6.5.1.2 Design

Each of the five conversation dyads recorded a conversation at assessment one and a second at assessment two. In the time period between the two conversations participants underwent assessment of language including picture naming, comprehension, reading, repetition and spelling (see study design, Chapter 5, section 5.1.1). Conversation one was analysed for stage one of the reliability study and both conversations were analysed for stage two.

TABLE 6.1 Background details of the five participants with aphasia and their conversation partners

Name	Person with aphasia				Conversation partner	
	Age	Sex	Years post-onset	Aphasia type	Sex	Relationship
HM	45	M	5	Broca's	F	Carer
PH	77	F	2	Anomic	F	Friend
SC	65	M	4	Mixed	F	Wife
BB	60	M	3	Anomic	F	Partner
DA	58	M	1	Broca's	F	Wife

6.5.1.3 Procedure

A conversation was tape-recorded by each couple in their own home or in a familiar environment (such as a day-centre). The tape recorder was left at the participant's home. Participants and their conversation partners were instructed to converse as normally as possible for at least ten minutes, and where possible for 15 minutes. No topic constraint was imposed, they were merely asked to talk as they would normally do. The researcher was not present during the recording, except in the case of PH where the researcher set up the tape-recorder for the two speakers, who were both unconfident about operating the equipment, and then left the room.

6.5.1.4 Transcription

The middle five minutes of the tape-recorded samples were transcribed, plus one minute either side to establish the context of any ambiguous speech. These extra two minutes were not analysed. Trained speech and language therapists who did not know the participants involved carried out the transcriptions in line with the descriptions given above (section 6.4). The conventions for transcription used here are those suggested by Levinson (1983). All phonological errors and neologistic output were phonetically transcribed in broad transcription.

6.5.1.5 Analysis

The raters listened independently to the tape recording, and studied the transcript, in order to clarify any uncertainties in the transcription. Each researcher then counted the instances of each variable outlined in the measure.

In order to analyse inter-rater reliability the resulting transcription was analysed independently by two researchers familiar with the measure. The author of this text analysed all the conversations and is referred to as rater one, and two other researchers analysed 50% of the conversations each and are referred to jointly as rater two). Thus for the five dyads described above the data consisted of one transcribed conversation of five minutes in length analysed by two separate researchers. Comparison of the data from the two raters constituted the inter-rater reliability analysis.

In order to analyse intra-rater reliability rater one analysed the specified data from conversation one on two separate occasions, at least one year apart. Comparison of the data from the two analyses constituted the intra-rater reliability analysis.

6.5.1.6 Item agreement

The measure developed here involves categorising behaviours and thus a degree of subjectivity is inevitably present. It was therefore essential that item by item agreement for the various events categorized under any given behavioural category was investigated. Clear definitions of the variables were developed but there is still some possibility of disagreement. Item agreement was investigated for all variables included in draft one.

There are two possible forms of analysis used to investigate item by item agreement. The first is widely used in similar studies (e.g. Nicholas and Brookshire, 1993; Brookshire and Nicholas, 1994; Doyle et al, 2000) and involves computing the proportion of agreed events from the total set of agreed and disagreed events. For example, rater one and rater two might each identify 10 instances of a given behaviour. They might agree upon eight of those instances, with a further four disagreements. This would give a proportion of agreement of $8/(4+8)$ or 67%. There is no accepted cut-off for proportional level of agreement in the published literature: in the analysis described here a minimal agreement of 60% was selected as the criterion.

This statistic fails to take account of agreement being reached by chance however. This is particularly the case where there is a closed set of behaviours and few categories, for example in labeling turns as substantive, minimal or other. A second statistical analysis used in assessing rater agreement in behavioural categorization studies is the Kappa statistic (Cohen, 1960). The essential difference between the method described above and the Kappa statistic is that the latter takes account of chance in its computation. The use of this statistic is recommended by various authors in the behavioural sciences and in particular within linguistics (e.g. Carletta, 1996). The level of Kappa deemed satisfactory varies across authors, but those recommended by Landis and Koch (1977) are used here and are shown in Table 6.2.

TABLE 6.2 Values of Kappa recommended by Landis & Koch (1977)

Level of agreement	Kappa value
Poor	< 0
Slight	0.00 – 0.20
Fair	0.21 – 0.40
Moderate	0.41 – 0.60
Substantial	0.61 – 0.80
Almost perfect	0.81 – 1.00

The possible levels of Kappa vary from -1.00 to $+1.00$. Gardner (1995) states that a Kappa value of 0.70 represents an acceptable level of agreement. The level of 0.70 is used as the criterion for satisfactory agreement in this study.

Where possible in the analyses presented here the Kappa statistic is used. Where this is not possible, i.e. where it is not possible to state the level at which a behaviour could be allocated to any given category by chance, the proportion of agreements is derived.

Details of the analyses are presented below.

6.5.1.7 Data and analysis

As there are variations in the number of instances of any given behaviour, with some behaviours producing many exemplars such as turns and speech units, and some producing few examples, for example, semantic errors, neologisms, the amount of data considered for a given variable differed. Within any given variable the same percentage of data was however analysed for both inter-rater and intra-rater reliability. The complete set of data for the five dyads, as analysed by the two raters at assessment one is given in Appendix Three.

For speech units and for turns and types of turns 20% of the total data from the five conversations was used. This amounted to one minute per dyad. For all other variables including topics, lexical retrieval, trouble indicating behaviours, and repair, the total conversation from each dyad was analysed.

Topics identified by rater one and by rater two were compared and the proportion of these that were agreed by both raters identified. Speech units, turns, trouble indicating behaviours and instances of repair were analysed in the same way. Types of turns were analysed additionally by considering both speakers' analyses, and using three codes for each turn: substantive turn, minimal turn, or other turn. Chance was therefore 0.33.

Lexical retrieval was analysed by coding each speech unit as a content word, a noun, a trouble indicating behaviour, or other. This gave four possible categories and chance was 0.25. As for turn types it is possible that poor agreement in one category could be masked by good agreement for other variables, and therefore individual proportional agreement for content words and nouns was also computed. For content words this involved taking all items coded as such by either rater, and computing the number of agreements and disagreements; similarly for nouns.

Within the class of trouble indicating behaviours there were eight possible categories. In order to establish item agreement within these categories these items were considered separately from other speech units, content words, and nouns. For example the total set of items coded as a phonological paraphasia by either rater were taken and the number of agreed items compared to the number of disagreed items. This procedure was repeated for all the trouble indicating behaviour categories.

The number of turns spent on repair was analysed by taking each turn identified by one or other rater as being involved in repair and computing agreement for each. Chance was thus 0.5. Repair types were analysed in two analyses. For successful and unsuccessful repair there were two categories, with chance at 0.5. For types of repair there were seven possible categories, so chance was lower at 0.14. Where chance is high (e.g. 0.5) the Kappa statistic takes this into account and demands a higher level of agreement than where chance is lower (e.g. 0.14).

6.5.2 Results and discussion of item agreement

Table 6.3 shows the proportion of agreement and the Kappa statistic where this is available for all variables considered in this stage of analysis for inter rater reliability and intra rater reliability.

There is no agreed acceptability level for proportion agreement in the published literature. In the figures shown in Table 6.3 values range from 0.29 to 0.99. A cut-off point of 0.60 for both inter rater and intra rater reliability was agreed upon. The cut off level of Kappa was 0.70.

Inter rater agreement for topics was 0.61 and intra rater agreement was 0.57 which is a moderate level of agreement for both. This reflects the difficulty in analysing sub-topics where a topic within a topic is introduced but which is rated as a new topic by only one rater. Topics were therefore excluded from the second draft of the measure.

TABLE 6.3 Item agreement for inter-rater and intra rater reliability

Variable	Interrater		Intra-rater	
	Proportion agreement	Kappa	Proportion agreement	Kappa
Topics covered	0.61		0.57	
Speech units	0.98		0.99	
Turn types	0.93	0.90	0.93	0.90
Substantive turns	0.98		0.98	
Minimal turns	0.73		0.73	
Lexical retrieval	0.86	0.79	0.90	0.87
Content words	0.77		0.78	
Nouns	0.84		0.93	
Trouble indicating behaviours:				
Semantic paraphasias/circumlocutions	0.42		0.34	
Phonological errors	0.68		0.81	
Neologisms	0.68		0.71	
Overuse of pronouns/proforms	0.29		0.42	
Comments	0.56		0.56	
Repetition of word or part word	0.56		0.54	
Pause of greater than two seconds	0.77		0.90	
Filled pauses	0.85		0.86	
Instances of repair	0.71		0.63	
Number of turns in repair (chance = 0.5)	0.78	0.56	0.80	0.60
Successful or unsuccessful repair (chance = 0.5)	0.86	0.72	0.89	0.78
Repair type (chance = 0.14)	0.29	0.17	0.53	0.45

Speech units were however reliably coded as such in inter and intra rater reliability.

This is not surprising as this analysis does not involve much judgement but merely involves counting items. There was also good agreement for turn types. The Kappa statistic for turn coding was high (0.90) for both inter rater and intra rater reliability.

Two independent raters and the same rater on two occasions are therefore able to reliably code turns into the categories using the definitions of the turns types provided in the measure. Individual agreement for turns types was high for substantive turns and

less so for minimal turns. Inspection of the minimal turn data revealed that disagreements were most likely where a person's turn could be interpreted either as a response to a question or as a minimal contribution. It was not felt however that tightening the criteria for inclusion into this category would help and that a degree of subjectivity was inevitable.

Agreement over categorization of items into one of the four categories of lexical retrieval (content word, noun, trouble indicating behaviour or other) was good when Kappa values were considered. This means that items are reliably coded into those four categories including the broad category of trouble indicating behaviours. Proportion agreement for content words was lower than that for nouns, but was above the cut-off of 60%. Disagreements were due mainly to failure by one rater or one rating occasion to code the verb 'to be' as a content word when it appeared in the third person singular. Agreement for nouns was high.

A number of trouble indicating behaviours showed poor agreement. These included: semantic paraphasias and circumlocutions, overuse of pronouns, comments, and repetition. The proportion of agreement for these variables fell below 0.70 in all cases and in some cases was far below this value. All were therefore eliminated from the second draft of the measure. Inspection of the data showed that semantic errors may arise and not be corrected thus may go unnoticed unless through a subsequent turn the error becomes clear. Overuse of pronouns is difficult to rate as there is a lot of unspecified referencing in normal conversation. Comments were agreed if they were an overt aside (e.g. *oh I can't think of the word*) but were less well agreed if they were less overt (e.g. *whatsername*). Repetition showed good agreement when a person produced the same word in succession but when other minimal speech units appeared between repeated instances of the first item agreement was poor. Pauses and filled pauses were reliably coded however and were included in the second draft.

The number of repairs in any given conversation, the number of turns spent on repair, whether a repair was successful or not, and the type of repair were considered here. Instances of repair showed good inter-rater but less good intra-rater reliability. The number of turns spent in repair was also satisfactory. Whether a repair was successful or not also showed good reliability. Repair types showed poor reliability. This demonstrates that two independent raters and the same rater on separate occasions were

able to reliably code repairs as to whether they were successful or not, but beyond that could not agree as to the form the repair took. One issue here concerns the amount of data. There were very few instances in some of the repair type categories, particularly in the unsuccessful repairs, as there were relatively few instances of repair in comparison with the number of repair categories. It may be that with a larger sample better agreement could be obtained. Repair types were nevertheless excluded from draft two.

6.5.3 Variables for draft two

As a result of the above analyses the variables considered to show satisfactory inter rater and intra-rater reliability were identified and these were then taken forward to the next phase of the reliability study, where test-retest stability was analysed. The total set of variables considered under this next phase (conversation measure draft two) is given in Appendix Four.

6.6 TEST-RETEST STABILITY

6.6.1 Participants

The participants for this part of the study were ten people with aphasia (five of whose data from conversation one had already been considered in the analysis of inter-rater reliability), and their chosen conversation partners. Details of the ten participants are given in Table 6.4.

6.6.2 Procedure

The data was collected in the same way as described for the inter-rater reliability study. The data under consideration here consists of the two conversations tape-recorded by each dyad at assessments one and two, at least eight weeks apart. In the intervening weeks no therapy took place although the participants with aphasia underwent language assessment (see Chapter Five, section 5.1.1). Each conversation was then transcribed, checked for accuracy of the transcription, and analysed using Draft Two of the measure (see Appendix Four). The data consist of values for all the variables within the revised measure for all ten participants, at two times of testing: assessment one and assessment two. The rater in question for all conversations was the author of this report. Thus the

data considered for the first stage of the analysis is included here, along with a second conversation for those five dyads (HM, PH, SC, BB and DA) and two conversations from the further five dyads (NK, IK, OL, KR, QP).

TABLE 6.4 Details of the ten participants with aphasia and their conversation partners for test-retest stability study

Name	Person with aphasia				Conversation partner	
	Age	Sex	Years post Onset	Aphasia type	Sex	Relationship
HM	45	M	5	Broca's	F	Carer
PH	77	F	2	Anomic	F	Friend
SC	65	M	4	Mixed	F	Wife
BB	60	M	3	Anomic	F	Partner
DA	58	M	1	Broca's	F	Wife
NK	52	M	4	Anomic	F	Wife
IK	68	M	4	Broca's	F	Wife
OL	65	F	2	Anomic	F	Friend
KR	38	F	12	Broca's	M	Husband
QP	65	M	5	Broca's	M	Brother

6.6.3 Data analysis

Raw data from the two conversations was analysed in order to establish stability of the occurrence of the relevant behaviours across test times. It remains the case however that considerable variation in two conversations may occur which masks underlying consistencies. This is because one person may speak more or less, or may contribute in different ways, in two separate conversations, thus the raw data may show considerable numerical differences, but the proportion of the person's contribution to the conversation, and the extent to which they encounter difficulties in a conversation may be constant. In order to analyse these factors, in addition to the investigation of the raw scores proportional data was also considered for test retest stability. The proportional data was selected on the basis of those variables which showed greatest stability when raw scores were analysed, and were divided by relevant denominators. Simple correlations were used to look at the relationship between scores at assessments one and two.

6.6.4 Results of the test-retest stability analysis

The raw data for all variables considered here are given in Appendix Five. The values of Pearson's R for the set of variables in Draft two of the measure are shown in Table 6.5

TABLE 6.5 Values of Pearson's R for all conversation variables at assessments one and two

	Variable	Pearson's R
B	SPEECH UNITS	
B1	Person with aphasia	0.872**
B2	Conversational partner	0.619*
C	TURN TAKING	
C1	Number of turns: PA	0.637*
C2	Number of turns: CP	0.634*
C3	Number of substantive turns: PA	0.733**
C4	Number of substantive turns: CP	0.789**
C5	Number of minimal turns: PA	0.745**
C6	Number of minimal turns: CP	0.793**
D	LEXICAL RETRIEVAL	
D1	Total content words	0.835**
D2	Total nouns	0.751**
D4	Number of phonological errors	0.362
D5	Number of neologisms	0.040
D10	Within turn pauses	0.128
D11	Filled pauses	0.828**
E	REPAIR	
E1	Instances of collaborative repair	0.606*
E2	Number of turns spent on repair	0.508
E3	Number of successful repairs	0.523
E4	Number of unsuccessful repairs	0.078

Critical values of R (df = 8) for a one-tail test are: R = 0.549, p = 0.05; R = 0.632, p = 0.025; R = 0.716, p = 0.01. * denotes variables significant at p = 0.05, ** denotes variables significant at p = 0.01.

R was significant at $p \leq 0.01$ for the person with aphasia's speech units, substantive turns, minimal turns, content words and nouns, and filled pauses, and for the conversation partner's substantive and minimal turns. R was significant at $p \leq 0.05$ for the total turns for both speakers, the conversation partner's speech units, and for instances of collaborative repair. These twelve variables were retained in the third draft of the measure.

At this stage only raw data has been considered. There is a strong argument however for using proportional data to look at outcomes (see Nicholas and Brookshire, 1993;

Perkins et al, 1999). In these data the variables speech units and number of turns were used as denominators.

6.6.5 Proportional data

Proportional data for those variables which showed good agreement in the first analysis entered this part of the analysis. These were: substantive turns, minimal turns, content words, nouns, filled pauses, and number of repairs. Only the person with aphasia's data is considered as it is here that change might be expected after therapy. For types of turns, the denominator was the total turns produced by the person, thus the proportion of substantive turns was derived by dividing the total number of substantive turns produced by the person with aphasia, by the total turns they produced. Similarly for minimal turns. For lexical retrieval the denominator used was the total number of speech units the person with aphasia produced, thus the proportion of content words the person produced was derived by dividing the total number of content words, by the total number of speech units they produced. Similarly for nouns and filled pauses. For repair the denominator was the total number of turns produced by *both* speakers in the conversation. The values relating to these variables are shown in Table 6.6.

The proportion of substantive turns, and the proportion of content words produced showed good stability in terms of the correlation. The proportion of nouns, the proportion of filled pauses, and the proportion of repairs showed adequate stability with r significant at $p = 0.05$. The proportion of minimal turns produced by the person with aphasia showed poor stability although minimal turns themselves showed adequate stability (see Table 6.5).

TABLE 6.6 Test-retest stability: proportional data

F	PROPORTIONAL VARIABLES	Pearson's R
F1	Substantive turns / turns	0.833**
F2	Minimal turns / turns	0.405
F3	Content words / speech units	0.884**
F4	Nouns / speech units	0.630*
F5	Filled pauses / speech units	0.712*
F6	Number of repair/turns	0.594*

Critical values of r ($df = 8$) for a one-tail test are: $R = 0.549$, $p = 0.05$; $R = 0.632$, $p = 0.025$; $R = 0.716$, $p = 0.01$. * denotes variables significant at $p \leq 0.05$, ** denotes variables significant at $p \leq 0.01$.

6.6.6 Discussion of test retest stability

The analysis of the test-retest stability data is not straightforward. This is to be expected, as there are elements of conversation which are likely to vary across occasions. From the test-retest data given above it is evident that certain elements are more stable than others. The amount of speech produced by the person with aphasia (speech units) shows good stability, demonstrating that speakers in the ten dyads tend to contribute a comparable amount of speech in separate conversations. This is less so with the speech units produced by the conversation partner.

The turns produced by the two speakers show adequate stability and the types of turns produced by the two speakers show higher levels of stability. When proportional data is considered the proportion of substantive turns is stable across two conversations, but the proportion of minimal turns is not. As the number of minimal turns was stable (Table 6.5), the low value of R in the proportional data must reflect the variation found in the denominator, the total number of turns the speaker produced.

The number of content words and the number of nouns produced by the person with aphasia are stable across occasions in both the raw data and the proportional data. Within the category of error productions a less satisfactory story emerges. Poor stability was found for the occurrence of phonological errors, neologisms, and pauses of greater than two seconds. Filled pauses showed satisfactory stability. Four aspects of repair were considered and one, the number of instances of repair, showed satisfactory stability. This was found in the raw data and in the proportional data.

For some of the data the small numbers of each behaviour may lead to the apparent extreme variability across test occasions, for example PH produced four neologisms in one conversation and nine in another. This is due to the types of targets being attempted in the two conversations. The issue of how long the sample must be is pertinent here. Where small numbers of each error type are produced in each conversation, the type of error being dependent upon the conversational context and to some extent the behaviour of the conversation partner, poor stability will emerge.

6.7 SUMMARY

In this chapter the rationale and method behind the development of a quantitative measure of lexical retrieval in conversation has been described. Two analyses were conducted to identify (i) which variables in conversation show inter-rater reliability and intra-rater reliability, and then (ii) which of those variables show stability over two separate conversations. Of those variables which entered the analysis a subset showed satisfactory levels in both analyses and these make up the final measure. The final set are shown in Appendix Six.

A note of caution is introduced here. Although numerical agreement is present for those which are included in the final measure, the results relate only to the data collected for this study. Conversation being an unconstrained activity, it is possible that future participants will not show satisfactory levels of stability on the variables outlined here. When the measure is used as an outcome measure for therapy it is therefore recommended that for any given individual two conversations should be analysed prior to therapy to establish that the variables in question show the necessary stability. This is also the case for the three participants who are the focus of this study (PH, SC, and KR).

CHAPTER SEVEN: ASSESSMENT RESULTS FOR PH

7.0 INTRODUCTION

In this chapter assessment and therapy results for PH are presented. Some personal background information is provided, followed by an analysis of tests of language processing. For details of the design and the methods used please refer to Chapter Five.

7.0.1 Background details

PH is a 77 year old woman who attended formal education up to the age of 14. Throughout most of her adult life she was a housewife, bringing up three children. She has lived in south London for her entire life. PH is right handed and, despite slight weakness in that hand as a result of the stroke, continues to use it for writing. Her hearing is within normal limits, and her vision is aided by glasses.

In 1995 she sustained a left hemisphere CVA which left her with aphasia, and a mild right-sided hemiparesis. No CT data is available. She commenced participation in this study in 1999, four years post onset. She attended speech and language therapy immediately after the stroke, and continued to receive review appointments but no direct therapy during her involvement in this study. PH has continued to live alone since the stroke with support with shopping and housework. She attends social activities regularly within her sheltered accommodation, but reported that she lacks regular sympathetic conversation partners.

At the time of her entry into this study PH presented with a predominately expressive aphasia, with fluent speech and a marked anomia. She is able to read and understand single words and simple sentences, thus is able to manage her own medical appointments, and order shopping through an item selection sheet which she gives to her home-help.

The following language sample demonstrates some of the difficulties PH has in connected speech.

PH: Cinderella narrative

she says they're going away for the night . she still don't look very happy . then she sits there and the witch . not the witch what's the other one . well she going around she want it to be all nice she got on her feet not ordinary people clothes . she looks out and there are two mouse and the big thing

Lexical retrieval difficulties are apparent in her production of 'witch' for fairy godmother, and her rejection of this shows awareness of the error. She attempts to explain slippers with a lengthy circumlocution containing a further semantic error: 'she got on her feet not ordinary people clothes'. In addition PH makes an error in irregular plural noun phrase production with 'two mouse' instead of mice.

7.1 PICTURE NAMING

7.1.1 Scores at baseline naming

The assessment used was a set of 200 black and white line drawings. Participants were asked to find the best single word to describe what they saw in the picture, and their last response within five seconds of seeing the picture was scored. For details of this set see Chapter Five (section 5.2). PH's scores at assessments one and two are shown in Table 7.1.

TABLE 7.1 PH's scores in picture naming at assessments one and two

Assessment One		Assessment Two	
Raw score (n=200)	%	Raw score (n=200)	%
65	32.5	75	37.5

7.1.2 Error analysis

Details of the responses made by PH in picture naming are shown in Table 7.2. The pattern of errors differs across the two assessments. PH got more items correct in the second assessment, although this was not significant (McNemar chi square = 1.397, df =

199, n.s.). The pattern of her responses altered with PH producing proportionately more semantically related responses in the second assessment and proportionately fewer

TABLE 7.2 PH's Picture naming responses in assessments one and two

Response	Assessment One		Assessment Two	
	Raw score	%	Raw score	%
Correct	65	32.5	75	37.5
Visual error	3	1.5	1	0.5
Semantic error	39	19.5	89	44.5
Phonological error	13	6.5	6	3.0
Non-words	14	7.0	9	4.5
Rejections	11	5.5	3	1.5
No response	54	27.0	17	8.5
Other	1*	0.5	0	0
TOTAL	200	100	200	100

*The other response in Assessment One involved PH attempting to write the word and achieving part of this.

no responses. It is possible that PH was prepared to attempt a response more readily on the second occasion (which occurred eight weeks after the first administration) as she had become used to the researcher and the test scenario. Examples of PH's errors are given in Table 7.3.

TABLE 7.3 Examples of PH's naming errors

Error type	Target	Response
Visual error	library	quiet in the
Semantic error	saxophone	trumpet
Phonological error	button	/bʌtəz/
Non-words	grave	/tətə/
Rejections	mermaid	not the submarine

The visual errors PH produced all involved naming a part of the picture. In the example given she read aloud a sign shown in the picture of the library. Semantic errors were numerous and of a variety of types. Most of the non-word responses PH produced were single syllables involving a consonant and a vowel and were possibly the start of an attempt at a word. There were no instances of more complex neologisms. Examples of the different semantic error types are shown in Table 7.4.

TABLE 7.4 Examples of PH's semantic errors

Error type	Target	Response
Semantic error: other	box	four
Superordinate	cabbage	vegetables
Semantic co-ordinate	factory	warehouse
Semantic subordinate	money	pound
Semantic associate	arrow	archer
Circumlocution	aerial	on top the t.v.
Semantically and visually related	boots	shoes
Semantically and phonologically related	penguin	pelican

The majority of PH's semantic errors were circumlocutions. The distribution of the error types is shown in Table 7.5.

The distribution of error types is similar across the two assessments, with a majority of circumlocutions in both cases. In these instances PH knew the target item and attempted to describe something about it, indicating retained access to semantic representations, but difficulty accessing phonological forms. There are relatively few examples of single lexical errors: a total of 18 in assessment one and 26 in assessment two.

7.1.3 Psycholinguistic variables influencing PH's naming

Early analyses of aphasic naming and the variables affecting performance concentrated on the effect of frequency on picture naming. Rochford and Williams (1965) found an effect of frequency in the people with aphasia whom they tested. Since then a number of studies have questioned the universality of this finding (see in particular Nickels and Howard, 1994). It is possible that the apparent frequency effect found by Rochford and Williams (1965) was due to the influence of other variables, a number of which co-vary with frequency (e.g. familiarity, age of acquisition). In attempting to identify the specific independent effects of the variables likely to influence naming the issue of covariance presents researchers with problems: concreteness and imageability covary, as do a number of lexical variables. It is important therefore that evidence for the influence of variables be taken from at least two sources. Factorial analysis using matched subsets has been used in a number of studies. More recently multiple regression has been applied to this area. As Ellis, Lum, and Lambon-Ralph (1996) report neither form of analysis will offer the true picture, but consideration of both forms may lead to a more balanced conclusion.

TABLE 7.5: PH distribution of semantic error types in assessments one and two

Semantic error type	A1		A2	
	Raw score	%	Raw score	%
Semantic error: other	2	5.1	5	5.6
Superordinate	0	0	3	3.3
Semantic co-ordinate	3	7.7	8	9.0
Semantic subordinate	0	0	5	5.6
Semantic associate	12	30.8	2	2.2
Circumlocution	21	53.8	63	70.8
Semantically and visually related	0	0	2	2.2
Semantically and phonologically related	1	2.6	1	1.1
TOTAL SEMANTIC ERRORS	39		89	

Both analyses were conducted here. The set of 200 pictures contained matched subsets for a number of variables whose influence could thus be analysed independently: imageability, animacy, operativity, familiarity, familiarity and frequency combined, age of acquisition, and length (number of syllables). In each analysis a number of items had been assigned to the two matched subsets, ranging from 30 items per set to 55 per set, thus the majority of the data were excluded from each individual analysis. This form of analysis is therefore weakened by the number of exclusions.

The second analysis used multiple regression to identify the contribution of the variables to PH's performance. A critique of and examples of applications of these techniques to single case data are described by Ellis, Lum, and Lambon-Ralph (1996).

7.1.3.1 Matched subsets

The results of the analyses for the matched subsets are shown in Table 7.6. The data shown is the result of analysing the two assessment times combined. This measure was used as it involved more data and was therefore more powerful.

The raw data for all variables showing a significant effect in any of the three analyses are shown in Table 7.7. This table shows numbers correct at assessments one and two only. The analyses show that there is a significant effect of imageability when the two sets are combined. The raw data show that high imageability items are named more

successfully than matched low imageability items, with roughly twice as many high as low being correctly named. There is an effect of animacy. The raw data show that there

TABLE 7.6 PH analysis of psycholinguistic variables in matched subsets

Variable	A1 plus A2	
	z score	p value
Imageability	2.81	0.003**
Animacy	2.00	0.02*
Operativity	2.06	0.02*
Familiarity	1.40	0.08
Familiarity and frequency	-1.04	0.15
Age of acquisition	-2.19	0.01*
Length	0.52	0.30

Table 7.6 shows z-scores and values of p for one-tailed tests. * $p \leq 0.05$, ** $p \leq 0.01$.

is an advantage for animate over inanimate items³². Operativity was significant with highly operable items named more easily than those with low operativity. Finally age of

TABLE 7.7 PH raw data numbers correct in matched subsets

Variable			Assessment one		Assessment two	
		n =	Score	%	Score	%
Imageability	High	55	20	0.36	27	0.49
	Low	55	11	0.20	14	0.25
Animacy	Animate	36	20	0.56	22	0.61
	Inanimate	36	12	0.33	16	0.44
Operativity	High	43	16	0.37	24	0.56
	Low	43	8	0.19	15	0.35
Age of acquisition	Early	40	13	0.33	18	0.45
	Late	40	7	0.18	10	0.25

³² PH shows an advantage for animate items. The more commonly reported finding is that of an advantage for non-living over living items usually in the context of herpes simplex encephalitis. There a few reported cases of people with aphasia who present with a category specific semantic impairment with an advantage for animate items. Further data relating to this phenomenon in PH has been collected by Dr Wendy Best and Astrid Schroeder.

acquisition was significant. The data show that there is an advantage for early acquired words over late acquired words and that this effect is present at both assessments.

This analysis has indicated that there are a number of variables which affect PH's naming success. Most of these are semantic in origin: imageability, operativity and animacy. Age of acquisition, which is thought to affect lexical access (see section 1.4.2.2), also has some effect on PH's naming. There is no evidence of any post-lexical factors affecting PH's naming success.

This analysis did not consider the following variables: concreteness (which correlates highly with imageability) and frequency (which correlates highly with familiarity). In the following analysis, regression techniques were used to identify the influence of the variables included in the above analyses plus concreteness and frequency.

7.1.3.2 Regression analysis

Of the 200 items included in the naming set, three had missing values for one or more variables. These items were excluded from the analysis leaving a total of 197 items. Further analysis revealed one significant multivariate outlier and this was removed from the data set leaving a total of 196 items³³. A further eight items were excluded on the grounds that they were difficult to categorise with regard to animacy (e.g. fairy, devil). This left a total of 188 items. This set was used for the analysis of the naming data for all three participants.

An inter-correlation matrix was computed to identify co-variance in the variables involved, and the relationship between each variable and the dependent variable. The dependent variable is the number of times correct in two administrations of the naming test. This is shown in Table 7.8.

The variables which are significantly associated with PH's naming are: imageability, concreteness, animacy and age of acquisition. As was found for the analysis using matched sets there was no effect of familiarity, frequency or word length. Unlike the finding for the matched subsets there was no significant relationship between naming and operativity. The strongest relationship was found between concreteness and

³³ Outliers were identified in line with Tabachnik and Fidell's (2001) recommendations

naming. Because of its strong relationship to imageability no matched sets analysis of this variable had been conducted.

**TABLE 7.8 Correlation matrix for psycholinguistic variables
and PH naming data**

	IMAG	CONC	OPER	ANIM	FAM	FREQ	AOA	SYLL	PHON
IMAG	1.00	-.360**	.073	.101	.141	.034	-.131	.213**	.231**
CONC		1.00	.356**	.137	-.158*	.007	.274**	.034	.059
OPER			1.00	.278**	.325**	-.047	-.201**	-.013	-.065
ANIM				1.00	.197**	.113	.152*	-.063	-.041
FAM					1.00	.672**	-.509**	-.114	-.178
FREQ						1.00	-.406**	-.081	-.195**
AOA							1.00	.418**	.448**
SYLLS								1.00	.843**
PHON									1.00
PH NAMING	.157*	-.333**	.116	-.173*	.066	-.027	-.235**	-.034	-.094

IMAG = imageability; CONC = concreteness; OPER = operativity; ANIM = animacy; FAM = familiarity; FREQ = frequency; AOA = age of acquisition; SYLL = number of syllables; PHON = number of phonemes. ** = correlation is significant at the 0.01 level (2 tailed); * = correlation is significant at the 0.05 level (2 tailed).

The general goal of a regression analysis is the identification of the fewest variables to predict the outcome on the dependent variable (Tabachnick and Fidell, 2001). The regression will be best when all the IV's are strongly correlated with the DV and not with each other. For the four IV's which have a significant correlation with the DV the values of R are relatively low: the highest value of R is for concreteness (0.333). It is therefore unlikely that a strong regression model will emerge. The relationship between the IV's showed some significant correlations, notably between imageability and concreteness. Tabachnik and Fidell (2001) advise that when variables are highly correlated (and they cite a value of R of 0.90 as an example of a high correlation) this implies redundancy in the model: two variables are present which measure more or less the same thing. In the data in Table 7.8 there are no such high correlations.

For theoretical reasons however it appears sensible to include only one of these two semantic variables in the equation. Distinguishing between the effects of these two is not interesting: what is interesting is the further evidence of a semantic variable operating on PH's naming. There is evidence from the matched sets analysis of an effect of imageability on PH's naming. Correlations show that concreteness has a stronger

relationship with naming however, so it is sensible to include this variable and not imageability in the model.

Animacy was shown to affect naming performance in the matched sets analysis. It showed an association with naming in the correlation analysis and therefore was included in the regression model. Likewise for age of acquisition. Operativity was not included in the model as there was no evidence of a relationship between this variable and naming. Operativity shows a correlation with concreteness and it may be that concreteness was a confounding variable in the matched sets analysis. On theoretical grounds it might be argued that a post lexical variable should be included. There is absolutely no evidence from either the matched sets analysis or from the correlation matrix that either number of phonemes or number of syllables affects naming however, and therefore no such variable was included. In analysing the data the same dependent variable was used throughout: the number of times an item was correctly named on two separate administrations.

Multiple regression was therefore used which looked at performance in picture naming as the dependent variable, and concreteness, animacy and age of acquisition as the independent variables. R for regression was significantly different from zero ($F = 10.467$, $df = 3, 184$, $p < 0.001$) with overall R^2 at 0.146. Two of the independent variables contributed significantly to the prediction of naming performance: concreteness ($t = 3.91$, $p < 0.001$; $sr^2 = 0.07$) and age of acquisition ($t = 1.982$, $p < 0.05$; $sr^2 = 0.02$). The three independent variables contributed another 0.05 in shared variability. In total 15% of the variability in naming performance was predicted by scores on these three independent variables.

The regression model corroborates to some extent the findings from the matched sets. There is a significant effect of a semantic variable on PH's naming. The matched subsets indicated that this was imageability, but the regression analysis points to concreteness as being the stronger predictor. Age of acquisition was a statistically significant factor in both the matched subsets and the regression. With concreteness and age of acquisition present there is no effect of animacy in the regression model ($sr^2 = 0.01$) indicating that the effect is partially explicable by the other variables: animacy showed significant correlations with operativity, familiarity and age of acquisition.

7.1.3.3 Discussion of predictor variables

The two methods used here have been used in previous studies of aphasic language. There are clear correspondences between the two methods with both showing effects of a semantic variable (imageability or concreteness) and age of acquisition. Imageability, operativity, animacy and age of acquisition showed effects in the matched subsets. Concreteness, imageability, animacy and age of acquisition showed a significant relationship with naming in the correlation. The multiple regression analysis revealed a strong effect of concreteness. This analysis also showed an effect of age of acquisition. In summary the evidence points to PH's naming performance being affected by these two variables, and that either of the statistical techniques used here may lead to other variables which correlate with these two showing an apparent effect. Without support from both forms of analysis these apparent effects must be viewed with suspicion.

7.1.4 Written picture naming

7.1.4.1 Method and Results

PH's written picture naming was investigated using a set of animate and inanimate items, matched for frequency, familiarity and age of acquisition. As PH presented with a possible advantage for animate over inanimate items in spoken picture naming, further investigation of this had been deemed important³⁴. The set contained 72 items: 36 animate and 36 inanimate. PH's performance in terms of number correct and her error pattern are shown in Table 7.9.

7.1.4.2 Interpretation of results

The results show that PH is able to access orthographic forms more readily than she can phonological forms (her spoken picture naming scores at assessments one and two were 0.33 and 0.38 correct). She makes semantic errors in both modalities. There was no effect of animacy in this set, with naming of both animate and inanimate sets at 20/36

³⁴ Further investigation of this aspect of PH's language processing was carried out by Dr Wendy Best and Astrid Schroeder and the results are reported in Astrid Schroeder's unpublished MSc thesis, and in Schroeder Best & Herbert (2001).

TABLE 7.9 PH's written picture naming: % correct and error types

	Correct	Semantic error	Orthographic error	No response
Total set (n = 72)	0.64	0.17	0.06	0.14

correct. Although this looks interesting at first glance, this phenomenon may be due to the particular set in question, which was not controlled for certain key variables which are known to affect PH (notably concreteness and imageability). PH can nevertheless access orthography, and thus may be able to use orthographic cues to access spoken forms.

7.1.5 Phonological and orthographic cueing of picture naming

The procedure for cueing picture naming is outlined in Chapter Five, section 5.4. The results for PH are shown in Table 7.10.

TABLE 7.10 PH Cueing of picture naming

Condition	Phonological	Orthographic
Extra time (n = 12)	0	2
Single cue (n = 12)	9	10
Choice of cues (n = 12)	5	10

Chi square analyses of the phonological cueing, comparing the three conditions was significant (chi square = 14.26, df = 2, p = 0.001). The more effective method of cueing for PH was the single cue with the choice of cues being effective in only five of the 12 trials. A second analysis comparing choice with extra time was also significant however (chi square = 4.04, df = 1, p = 0.04).

The same analysis was conducted on the orthographic data. Again, there was a highly significant effect (chi square = 14.96, df = 2, p = 0.001). In this case no further analysis was conducted as both single cue and choice of cues were equally effective.

7.1.6 Interpretation of naming data

PH's attempts at naming a set of 200 words on two occasions, and a related study investigating her ability to respond to phonological and orthographic cues, reveal the following main findings. PH has a severe anomia with percentage scores of less than 40 on repeated tests of picture naming. Her errors are predominately circumlocutions, indicating that she has access to semantic representations, and failures to respond. Variables influencing her picture naming are semantic and lexical. She is able to respond to phonological and orthographic cues, and both a single cue and a choice of two cues were effective. It is hypothesised at this stage that PH has a possible semantic deficit, and a deficit affecting access to lexical representations. Her good response to cueing suggests lexical forms are available.

7.2 SEMANTIC PROCESSING

In order to investigate her semantic processing PH was tested on three tests of semantics: Pyramids and Palm Trees, and spoken and written word to picture matching.

7.2.1 Pyramids and Palm Trees

The three picture version of this test was used to investigate non-lexical semantic processing. PH scored 47/52 correct, which is outside the normal range of performance.

7.2.2 Word to picture matching

Spoken and written word to picture matching tests were administered. PH scored 29/30 correct in the written version, which she completed speedily and with no hesitation, and 28/30 correct in the spoken version, for which she requested repetition for three items. Both scores are within the normal range for the tests. PH's errors in the two tests all involved selection of the semantically related foil.

7.2.3 Interpretation of results

The data from picture naming suggested a possible semantic deficit as PH's naming was strongly predicted by a number of semantic variables. Counter to this theory, there were

relatively few single word semantic errors in her spoken picture naming, and she produced mainly circumlocutions describing items she could not name, indicating some semantic access. The results of tests of semantic processing indicate that if PH has a semantic deficit it is very mild: her performance is within normal limits on two tests, but she is outside the normal range for Pyramids and Palm Trees.

7.3 AUDITORY DISCRIMINATION

7.3.1 Test performance on minimal pairs

Auditory discrimination was tested using the Action for Dysphasic Adults non-word minimal pairs. PH scored 27/40 on this test which is significantly better than chance (Binomial test: $p = 0.019$).

7.3.2 Error analysis

Of the 13 errors ten were false negatives, where she called a same pair different, and three were false positives, where she called a different pair same. In other words, for the 'different' pairs, she scored 17/20 correct which is significantly better than chance (Binomial test: $p = 0.0013$), and for the 'same' pairs, she scored 10/20 correct, thus was at chance in the latter set. For the three false positives there was no pattern, in terms of phonetic distance between the two sounds, to the type of differences that caused the errors.

7.3.3 Interpretation of results

PH showed some difficulty with auditory analysis in this task. Her errors show that she was at chance in the 'same' set being unable to judge whether these pairs were the same or not. This may be due to: poor auditory acuity and a lack of confidence therefore in her own auditory perception when items were perceived as very similar; a misunderstanding of the task, where she expected more 'different' pairs; or a real effect of a deficit in auditory analysis. Her relatively good performance on the 'different' set suggest auditory analysis is better than the score of 27/40 would suggest. Further evidence to support the rejection of a hypothesis of impaired auditory analysis comes

from two sources: she did not select the phonologically related foils in spoken word to picture matching, and she showed a good response to spoken cues in picture naming.

7.4 SHORT TERM MEMORY

7.4.1 Test performance

This was tested with four different tests (see section 5.3.3). PH scored as follows on these tests: digit span 4.3, letter span 4.3, phoneme span 2.5, picture pointing span 3.5.

7.4.2 Interpretation of results

The data indicate that PH has some deficit in auditory short-term memory. Phoneme span is particularly poor. This involves the auditory perception of and repetition of strings of phonemes, for which one cannot conjure a visual image to aid retention. The other three tasks involve items for which one can use visual mnemonics (numbers, letters and pictures). In this sense phoneme span is a more pure assessment of auditory short term memory. PH may also have some deficit in auditory analysis. If this were the sole explanation of her difficulty with phoneme span however one would expect a worse score in letter span, which involves auditory perception of and repetition of consonant and vowel combinations with several neighbours (e.g. B, perceived as /bi/, C as /si/, D as /di/ and so on). It is perhaps a combination of these factors which produces the poor score in phoneme span: the lack of visual mnemonics, an auditory analysis deficit, and an auditory short term memory deficit.

7.5 READING ALOUD AND REPETITION

7.5.1 Reading aloud words

PH scored 177/182 correct (0.97). The five errors were on the following words: simile, prelude, protocol, camel and camera. She read aloud simile as 'similar', and prelude protocol, camel, and camera as phonologically related non words. There are few errors here so little can be made of the data. It is worth noting that three of the errors were on words from the low imageability / low frequency subset (simile, prelude and protocol). However, PH got the remaining 24 items in this subset correct: this included items such

as anecdote and paradox. A possible explanation is that PH did not know these three words. The other two errors on common words (camel and camera) are difficult to explain but occur in the context of the otherwise excellent access to output phonology from written stimuli.

7.5.2 Reading aloud non-words

PH scored a total of 9/26 correct when the criterion is accurate production of the whole phoneme string. She showed a length effect scoring 7/10 for one syllable items, 2/10 for two syllable items, and 0/6 for three syllable items (Jonckheere Trend Test: $z = 2.74$; one tailed $p = 0.003$).

Closer analysis reveals better performance than this score would suggest. She produced 22/26 correct initial phonemes, and when the total set of phonemes are considered she scored 101/130 (0.78) correct phonemes in the correct order.

7.5.3 Repetition of words

PH scored 176/182 correct (0.97). The errors were on the following words: realm, magazine, anecdote, coffer, prelude and grave. For the majority of these cases she appeared not to recognise the word (realm, anecdote, coffer, prelude and grave). For magazine she produced a phonemic error.

PH read aloud realm, anecdote coffer and grave without any difficulty and it is hypothesised that she does know these words. She was unable to read 'prelude' aloud however and it is probable that she does not know this word. It is possible that PH has a mild auditory analysis deficit which is insignificant when a word is well known to her. For low imageability low frequency words this may occasionally cause difficulties.

7.5.4 Repetition of non-words

PH scored a total of 15/26 correct when the criterion is accurate production of the whole phoneme string. She showed a length effect scoring 8/10 for one syllable items, 5/10 for two syllable items, and 2/6 for three syllable items (Jonckheere Trend Test: $z = 2.42$; one tailed $p = 0.0078$).

However, she produced 23/26 correct initial phonemes, and when the total set of phonemes are considered she scored 116/130 (0.89) correct phonemes in the correct order.

7.5.5 Interpretation of reading and repetition results

PH showed evidence of good access to phonological output representations in both reading aloud and repetition. There is little to be concluded from the errors she produced in these tasks as they are few, but there is some evidence of low imageability / low frequency items causing more problems. These data indicate the PH has intact phonological representations, and provide corroboration of the evidence from the effective use of cues.

Processing of non-words was less good. Reading was worse than repetition, indicating better trans-coding to phonology from auditory than from visual stimuli. Both the route from auditory analysis to spoken output and that from graphemic analysis to spoken output are therefore impaired, and indicate that PH is relying heavily on whole word lexical routes to read and repeat real words. Although there is impairment in both sub-lexical routes PH is nevertheless able to use them to some extent, shown by her good production of initial phonemes and the overall percentage of correct phonemes. Her difficulties in repetition may relate to the possible mild deficit in auditory analysis and her deficit in phoneme span. Her deficit in reading aloud suggests that the grapheme to phoneme route is more impaired than the auditory sub-lexical route. However, PH was equally able to respond to sub-lexical orthographic cues as to sub-lexical phonological cues.

7.6 TESTS OF INTERNAL PHONOLOGY

7.6.1 Test performance

On the picture homophone test (where the person has to identify the two items which have the same phonology, e.g. flower and flour), PH scored 13/20, which is significantly better than chance (Binomial test: $p = 0.003$). On a test of initial phoneme knowledge, where the person has to select the item which has the same initial phoneme

as the target item, PH scored 1/20. In this test all of PH's errors involved the selection of the semantically related foil.

7.6.2 Interpretation of results

These tests of internal phonology are notoriously difficult for people with aphasia. It has already been ascertained that PH has good phonological representations, that these are readily accessible via spoken and written input, and that they can be cued by spoken and written cues. In the tests reported here more overt meta-linguistic knowledge is probed. PH scored above chance in the homophone task but did not appear to grasp the initial phoneme task. PH does therefore appear to have some access to internal phonological knowledge. Evidence from her response to cues supports this contention. Given a choice of cues she is able to select the correct one in attempting to name a target word (section 7.1.5).

7.7 SENTENCE COMPREHENSION

7.7.1 Test performance

In spoken sentence comprehension PH scored 11/16 correct, making errors on two reversible active SVO sentences, and on three sentences with embedded clauses such as "The shoe under the pencil is red". In written sentence comprehension PH scored only 4/16 correct (= chance). Errors were made on all sentence types.

The discrepancy between auditory and written sentence comprehension is marked. Written comprehension was selected as a language control task however, and thus was re-administered at key points throughout the study. At the second administration of this task PH scored 11/16, and made errors on reversible SVO sentences and on sentences with embedded clauses. This exactly mirrored the performance seen in spoken comprehension.

7.7.2 Interpretation of results

PH's performance on the initial administration of the written sentence comprehension test appeared to identify a serious deficit in reading comprehension, at odds with the

performance seen in spoken comprehension. The second administration of this task showed a similar pattern across spoken and written input, and indicates that PH has a central grammatical impairment affecting decoding of meaning from word order information and from parsing of phrase structure (which is necessary to decode embedded phrases).

7.8 NON VERBAL PROCESSING

7.8.1 Non verbal problem solving

Her score on a subset of the Raven's Progressive Matrices was 11/12, showing that visual problem solving was relatively intact.

7.8.2 Visual processing

PH was able to bisect horizontal lines successfully indicating that there was no hemianopia. PH scored 24/32 on the BORB (Birmingham Object Recognition Battery: Riddoch and Humphreys, 1987) Object Decision A Hard task which is within normal limits (normal range 22-30 correct).

7.8.3 Interpretation of results

PH's good score on the Raven's Progressive Matrices indicates good visual problem-solving skills. Her good score on the BORB subtest indicates that her processing of visual stimuli is intact and thus semantic/visual errors such as those seen in her few errors in word to picture matching cannot be explained in terms of a visual processing deficit.

7.9 CONVERSATION

The data considered here are the two conversations produced at assessments one and two. These data also formed part of the analysis of the reliability of the conversation measure described in Chapter Six. In the discussion of PH's data, in particular the qualitative analyses, the issue of the reliability of subjective interpretation of the data needs to be borne in mind. Nevertheless it is possible to draw some tentative

conclusions about PH's language processing difficulties and how these affect her conversation.

7.9.1 Turn taking

PH's conversations with her friend revealed that she is able to take turns and contribute meaningfully to conversation. It is notable that the friend encouraged PH to speak throughout, and administered a number of cues for PH when she was in difficulty. These behaviours are unlike any seen in normal everyday conversation, where difficulties are solved quickly, and the act of drawing attention to a language processing problem would be seen as rude or insensitive. Thus, this contrived conversation differs from normal conversation, and must therefore be viewed as a special case. PH reported that in general conversation was very difficult for her. One possible reason is that the time pressures of turn-taking prohibited the satisfactory resolution of PH's word finding difficulties.

Conversation Sample 7.1

- 1 PH they were all out 'n and the girls /ə/ /ə/ people they had were out there
- 2 SP did they
- 3 PH yeah
- 4 SP lots of lovely flowers then
- 5 PH yeah [mmm
- 6 SP [mm
- 7 PH well they had one each and the had the what one whatsnames as well then
- 8 they come up er to the the the /fɑ/ er the father . one of them had erm erm it .
- 9 was one that you can /hɛl/ /hɛl/
- 10 SP oh yes
- 11 PH and so did the other one as well yeah . they didn't give anything for the girls
- 12 SP didn't they
- 13 PH thought they would have done . usually you do don't you
- 14 SP yes you give them gifts
- 15 PH yes but they didn't least not as far as I know

PH: person with aphasia

SP: conversation partner

Extract from PH's second conversation

In this extract, in which PH is describing a recent family wedding she attended, she shows that she is able to initiate turns (lines 7, 11, 13 and 15) and topics³⁵ (line 11) in conversation. A number of her turns are substantive (lines 1, 7, 11, 13 and 15) and there are minimal turns too (line 5).

7.9.2 Lexical retrieval

Although PH has a severe anomia (see picture naming data section 7.1) she is able to produce nouns in conversation (lines 1, 8, 11). She also uses some high frequency verbs (e.g. be, have, come, give) and occasional adjectives and adverbs.

7.9.3 Trouble indicating behaviours

PH shows evidence of a number of trouble indicating behaviours including semantic errors³⁶ (e.g. in line 1 PH said 'girls' then corrected this to 'people'). The predominant feature leading to breakdowns in PH's conversations is that of lexical retrieval. In sample 7.1 the following evidence of word finding difficulties is present: semantic paraphasia (line 1), phonological paraphasia (line 8: /fɑ/ father), comments (line 7: what'sname), filled pauses (line 8)³⁷.

7.9.4 Repair

Although PH experiences a number of difficulties in completing her turns, and shows evidence of a number of trouble indicating behaviours, there are relatively few instances of collaborative repair within the conversations (seven in each five minute sample). This is mainly because of an apparent tolerance on the part of her conversation partner of a lower level of mutual comprehension, or inter-subjectivity, than is normally the case. For example in Sample 7.1, lines 7 to 9 PH produces a long turn within which she encounters a number of problems (phonological paraphasia, comments, filled pauses) and at the conclusion of which the naïve observer is unsure what the intended topic and proposition are. PH's conversation partner's next turn (line 10) 'oh yes' indicates that she accepts the proposition of PH's turn and it signals agreement with or approval of the

³⁵ See Chapter Six, for a discussion of inter-rater and intra rater reliability for topics and turns

³⁶ Note that certain phenomena in conversation showed poor reliability (see Chapter Six). One such aspect was semantic errors where poor inter rater and intra rater reliability were found.

³⁷ The same caveat described with regard to semantic errors applies to these error types, apart from filled pauses.

turn. An appraisal of the content would demand an initiation of repair at this point, which the partner singularly fails to do. This serves to hand the floor back to PH to either continue that topic or to initiate a new topic. In this way the partner protects PH from having to fully engage with and solve her aphasic word finding difficulties and shows a sensitivity to the social consequences of failing in conversation. Although this graceful behaviour may save face, it may also engender a lack of ultimate satisfaction for both parties, given that there is thus a consequent lack of inter-subjectivity in the conversations. Some evidence to support this contention comes from the fact that although PH's friend offered to carry out the conversations in the study, PH reported that she rarely conversed with the friend outside of these occasions.

7.9.5 Conversation measure

The variables found to have good inter-rater and intra-rater reliability, and good test-retest reliability in the conversation measure (see Chapter Six) made up the final version of the measure. This was then used to analyse PH's data at assessments one and two³⁸. These data are shown in Table 7.11.

The analyses carried out with the conversation measure show that most of PH's turns are substantive (0.68 and 0.59 at assessments one and two respectively). She also produces some minimal turns (0.21 and 0.18 of her turns are minimal at assessments one and two respectively). She produces a large number of content words (163 and 141) making up around 0.25 of her spoken output. Noun production makes up a relatively small part of her content word production with the majority of her content words being verbs. This difficulty in noun production is in line with her performance on spoken naming tests (section 7.1).

In comparison with her conversation partner PH produces more speech units overall, although the number of turns and the proportion of substantive and minimal turns was comparable. This reveals that PH's turns are longer than her partner's. Given the artificial nature of the task this cannot be taken as proof that PH habitually produces more speech in conversation than does her partner. It is more likely that this reflects the partner's encouragement to PH throughout both to take turns and to complete turns when difficulties arise.

³⁸ The data here are from the analyses carried out by the author of this report.

TABLE 7.11 Conversation measure: data for PH at assessments one and two

	Variable	Assessment One	Assessment Two
B	Number of speech units		
B1	Person with aphasia	566	544
B2	Conversational partner	302	284
C	Turn taking		
C1	Number of turns: Person with aphasia	72	71
C2	Number of turns: Conversational partner	69	66
C3	Number of substantive turns: Person with aphasia	49	42
C4	Number of substantive turns: Conversational partner	46	39
C5	Number of minimal turns: Person with aphasia	15	13
C6	Number of minimal turns: Conversational partner	20	21
D	Word retrieval and speech errors (PA only)		
D1	Total number of content words (excluding paraphasias)	163	141
D2	Total number of nouns (subset of content words)	30	50
D11	Number of filled pauses within PA's turn	7	15
E	Repair		
E1	Total instances of repair	7	7
F	Proportional data		
F1	Substantive turns / turns (PA)	0.68	0.59
F2	Minimal turns / turns (PA)	0.21	0.18
F3	Content words / speech units	0.29	0.26
F4	Nouns / speech units	0.05	0.09
F5	Filled pauses / speech units	0.01	0.03
F6	Repair / total turns	0.05	0.05

7.10 SUMMARY OF PH'S LANGUAGE PROFILE

PH presents with fluent anomic spoken output. She is able to take part relatively effectively in conversation although there are frequent breakdowns due to word finding difficulties. She also has difficulties in accessing written word forms although this is less severe than her spoken word finding deficit. She has good comprehension for everyday conversation although testing reveals a deficit in deriving meaning from syntactic relationships.

Her severe anomia occurs in the context of otherwise relatively preserved single word processing, with good comprehension, reading aloud and repetition of single words. The deficit appears to arise from a semantic processing deficit which is relatively mild but is apparent from three lines of evidence, and a lexical access problem. The semantic deficit is evident from the fact that PH shows a deficit in one test of input processing (Pyramids and Palm Trees), she produces some co-ordinate errors in spoken output in picture naming, connected speech tasks, and in conversation (although see caveats regarding the reliability of the conversation data); and her picture naming is predicted by semantic variables (imageability, concreteness, and possibly animacy). Evidence for this being a mild deficit comes from the fact that her performance in verbal input tasks is within normal limits, and that she mainly produces semantic information about the target in spoken picture naming, with co-ordinate errors being more rare. This indicates that she can access semantic information but fails to access phonology.

PH's phonological representations appear to be intact and available as shown by her good performance in transcoding from input to spoken output when semantic mediation is not required, in reading aloud and repetition, and from the evidence of successful word retrieval when a phonological or orthographic cue is provided. Age of acquisition also predicts naming to some extent and therefore further supports the contention that there is an additional problem at the lexical level. It is therefore proposed that PH has a mild semantic deficit and a further deficit in mapping from semantic representations onto phonological forms.

CHAPTER EIGHT: ASSESSMENT RESULTS FOR SC

8.0 INTRODUCTION

In this chapter assessment results for SC are presented. Personal background information is provided, followed by an analysis of tests of language processing, and data from conversation. For details of the design and the methods used please refer to Chapter Five.

8.0.1 Background details

SC is a 65 year old man who sustained a left CVA in 1994, five years before becoming involved in this research. SC was educated up to age 18, and after leaving school he worked in a variety of jobs, most recently as a driver for a car-hire company. He was admitted to hospital for angioplasty following a cerebral haemorrhage, but sustained an ischaemic stroke while undergoing surgery. This left him with a right-sided hemiplegia, a right homonymous hemianopia, and aphasia. He presents now with a right-sided hemiparesis, and although he is mobile around the home he uses a wheelchair for outings. He was right handed but now uses his left hand for writing³⁹. Hearing was reported to be normal. SC presented with a progressive visual impairment, for which he had no definitive diagnosis although SC and his wife had been told this was a form of macular degeneration. Despite this impairment he was able to distinguish pictures and letters with the aid of glasses, but he found bright light intolerable. Care was taken throughout the data collection to compensate for this deficit and to ensure that all visual stimuli were adequately perceived, by using large text and enlarged picture materials.

Prior to the stroke SC was a good conversationalist with a penchant for gossip and anecdotes, and a gift for story telling which remains even in the context of the aphasia. He lives with his wife and has a large supportive family who visit frequently. He reported that he finds one to one conversation good but in a family gathering often struggles to keep up with the topic. He presents with expressive and receptive aphasia. His comprehension is adequate for simple conversation, but his difficulties in formulating sentences and finding words, and a deficit in self-monitoring skills, mean expression is seriously compromised.

³⁹ SC does not write with a pen and paper but 'skywrites' letters to aid word retrieval

SC: Cinderella narrative

*and her run down she er dropped her /s/ slipper but still running like a lunatic on time .
onto the /kɒst/ and her /hɔ/ horse charging out quickly to where her's got to be
when she got there whoof everything back to normal
and the prince said I saw this girl and I am going to marry her*

SC's spoken output is fluent with frequent word finding difficulties and paraphasias, and he has sentence construction difficulties. The extract above shows some of these difficulties. SC is able to bypass his severe anomia in connected speech by the use of relatively empty circumlocutions for example 'to where her's got to be' which has no obvious referent. He also produces phonological errors, and neologisms.

8.1 PICTURE NAMING

8.1.1 Scores at baseline naming

The assessment used was a set of 200 black and white line drawings. For details of this set see Chapter Five (section 5.2). SC's scores at assessments one and two are shown in Table 8.1.

TABLE 8.1 SC's scores in picture naming at assessments one and two

Assessment One		Assessment Two	
Raw score (n=200)	%	Raw score (n=200)	%
67	33.5	74	37.0

8.1.2 Error analysis

Details of the responses made by SC in picture naming are shown in Table 8.2.

SC's performance in terms of number correct and the distribution of error types is comparable across the two test times. Like PH (when one considers her profile across the two test times) SC makes mainly semantic errors and no responses.

TABLE 8.2 SC's Picture naming responses in assessments one and two

Response	Assessment One		Assessment Two	
	Raw score	%	Raw score	%
Correct	67	33.5	74	37.0
Visual error	7	3.5	2	1.0
Semantic error	37	18.5	36	18.0
Phonological error	3	1.5	5	2.5
Non-words	3	1.5	15	7.5
Rejections	7	3.5	9	4.5
No response	76	38.0	59	29.5
Other	0	0	0	0
TOTAL	200	100	200	100

Examples of SC's errors are given in Table 8.3.

TABLE 8.3: Examples of SC's naming errors

Error type	Target	Response
Visual error	stable	door
Semantic error	magazine	book
Phonological error	elephant	/ɛfəlɒnt/
Non-words	kennel	/tʃuhænd/
Rejections	camel	cow no

Examples of the different semantic error types are shown in Table 8.4. The distribution of the semantic error types is shown in Table 8.5.

Again there is consistency between the two test times with the distribution of error types being similar in both assessment one and assessment two. The majority of semantic errors are semantic associates or circumlocutions at both assessment times. There are few examples of lexical access errors such as superordinate, co-ordinate or subordinate errors. The semantic associate errors can be interpreted in one of two ways: they either imply that SC has good access to semantics and, unable to access phonology, provides another associated word, or they imply that SC has poor semantic access and is able to identify only a broad semantic field associated with the target. The circumlocutions indicate that SC has good access to the semantic representation but fails to access the phonology. He then is able to describe the target using other words. These circumlocutions were very specific (such as 'a bird mouth' for target beak) and identified the target unequivocally. This is not always the case with productions rated as

TABLE 8.4 Examples of SC's semantic errors

Error type	Target	Response
Semantic error: other	footballer	player
Superordinate	pineapple	fruit
Semantic co-ordinate	ladder	stairs
Semantic subordinate	-	-
Semantic associate	aerial	television
Circumlocution	beak	a bird mouth
Semantically and visually related	garage	shed

circumlocutions⁴⁰. So far SC presents in a very similar way to PH with good semantic access, at least as shown by the circumlocutions. The data from the other error types, including the semantic associate errors, imply a possible semantic deficit.

TABLE 8.5 distribution of semantic error types in assessments one and two

Semantic error type	Assessment one		Assessment two	
	Raw score	%	Raw score	%
Semantic error: other	2	5.4	0	0
Superordinate	3	8.1	3	8.3
Semantic co-ordinate	0	0	3	8.3
Semantic subordinate	0	0	0	0
Semantic associate	13	35.1	13	36.1
Circumlocution	14	37.8	14	38.9
Semantically and visually related	4	10.8	2	5.6
Semantically and phonologically related	1	2.7	1	2.8
TOTAL SEMANTIC ERRORS	37		36	

8.1.3 Psycholinguistic variables influencing SC's naming

As for PH two analyses were conducted here. The first analysis involved the matched subsets contained within the larger set of 200 pictures. As stated in Chapter Seven (section 7.1.3) the variables for which there were matched sets are: imageability, animacy, operativity, familiarity, familiarity and frequency combined, age of acquisition, and length (number of syllables). The second analysis using multiple

⁴⁰ There is an issue of reliability here which is not addressed in this paper or indeed in most published papers. This concerns agreement over coding errors into categories. In order to ascertain whether there is consistency in raters' coding, inter-rater reliability needs to be analysed. This could be done by comparing two raters categorization of a subset of the data, then using Cohen's Kappa to analyse the results. This was not done here for reasons of time and availability of raters but is recommended for future studies.

regression techniques was also carried out to identify the contribution of the variables to performance.

8.1.3.1 Matched subsets

The results of the analyses for the matched subsets for SC are shown in Table 8.6. The data relate to the two combined assessment sets. This measure was used as it involved more data and was therefore more powerful.

TABLE 8.6 SC analysis of psycholinguistic variables in matched subsets

Variable	A1 plus A2	
	z score	p value
Imageability	-0.83	0.35
Animacy	-2.16	0.02*
Operativity	-0.06	0.48
Familiarity	1.11	0.13
Familiarity and frequency	2.45	0.007**
Age of acquisition	0.04	0.48
Length	1.17	0.12

Table 8.6 shows z-scores and values of p for one-tailed tests. * $p \leq 0.05$, ** $p \leq 0.01$.

SC shows a significant effect of animacy, naming more inanimate than animate items correctly. Interestingly the effect is the opposite to that usually found in people with aphasia where an advantage for living things is the usual pattern (e.g. Howard et al, 1997). This is also in contrast to PH who showed an advantage for living things.

Familiarity and frequency combined has a significant effect. As this includes two variables it may be that one is the underlying cause. The raw data for the two variables showing a significant effect are shown in Table 8.7. This table shows numbers correct at assessments one and two only.

The raw data further clarify the existence of an advantage for inanimate things over animate things, in both assessment one and assessment two. They also show a difference between assessments one and two in terms of the effect of familiarity and frequency combined: this shows little difference at assessment one and a marked difference at assessment two.

TABLE 8.7 SC numbers correct in matched subsets (raw data)

Variable		n =	Assessment one		Assessment two	
			Score	%	Score	%
Animacy	Animate	36	8	22	6	17
	Inanimate	36	17	47	13	36
Familiarity and frequency	High	50	17	34	26	52
	Low	50	13	26	11	22

In this instance where there is some ambivalence about the effect of these variables on naming, the alternative analysis using regression techniques is more than warranted.

The analysis of the matched subsets did not consider concreteness (which correlates highly with imageability) and frequency (which correlates highly with familiarity). In the following analysis, regression techniques were used to identify the influence of the variables included in the above analyses plus concreteness and frequency.

8.1.3.2 Regression analysis

The same set of 188 items that were analysed for PH were analysed here for SC (see section 7.1.3.2). An inter-correlation matrix was computed to identify co-variance in the variables involved, and the relationship between each variable and the dependent variable (see Table 8.8). The dependent variable is the number of times SC named an item correctly in the two administrations of the naming test. The final line gives the relationship between the variables and the dependent variable (SC's naming on two occasions).

The variables which are significantly associated with SC's naming are: animacy, familiarity, frequency, age of acquisition, and word length as measured by number of phonemes. These results agree with the findings from the matched sets analysis, with the exceptions of age of acquisition, and a possible length effect. The emergence of age of acquisition as a strong predictor variable is puzzling given that this variable exerted no influence when analysis was confined to matched subsets. There are however strong inter-correlations between age of acquisition and both familiarity and frequency. There are three reasons to doubt the existence of a length effect: this was not found in the matched sets analysis; number of phonemes correlates highly with age of acquisition ($R = 0.448^{**}$) and quite highly with word frequency ($R = -0.195^{*}$), thus this apparent effect may be due to other variables; and the alternative measure of word length,

TABLE 8.8 Inter-correlation matrix for the set of 188 items and for SC's naming

	IMAG	CONC	OPER	ANIM	FAM	FREQ	AOA	SYLL	PHON
IMAG	1.00	-.360**	.073	.101	.141	.034	-.131	.213**	.231**
CONC		1.00	.356**	.137	-.158*	.007	.274**	.034	.059
OPER			1.00	.278**	.325**	-.047	-.201**	-.013	-.065
ANIM				1.00	.197**	.113	.152*	-.063	-.041
FAM					1.00	.672**	-.509**	-.114	-.178
FREQ						1.00	-.406**	-.081	-.195**
AOA							1.00	.418**	.448**
SYLLS								1.00	.843**
PHON									1.00
SC	.094	.009	.060	.170*	.237**	.304**	-.246**	-.090	-.154*
NAMING									

Values are for Pearson's R. IMAG = imageability, CONC = concreteness, OPER = operativity, ANIM = animacy, AoA = age of acquisition, FAM = familiarity, FREQ = Celex log combined frequency, SYLLS = number of syllables, PHON = number of phonemes. Significance levels of R are ** $p \leq 0.01$, * $p \leq 0.05$.

syllable number, does not predict word retrieval. This variable was not therefore included in the regression analysis. Although there are significant correlations between the independent variables, most notably between familiarity, frequency and age of acquisition, it was decided to include all three in the regression. The correlations, whilst high, are not extremely so.

Multiple regression was used which looked at performance in picture naming as the dependent variable, and animacy, familiarity, frequency, and age of acquisition as the independent variables. R for regression was significantly different from zero ($F = 7.55$, $df = 4, 183$, $p < 0.001$) with overall R^2 at 0.142. Three of the independent variables contributed significantly to the prediction of naming performance: animacy ($t = 2.578$, $p = 0.011$; $sr^2 = 0.03$), frequency ($t = 2.64$, $p = 0.009$; $sr^2 = 0.03$) and age of acquisition ($t = 2.545$, $p = 0.012$; $sr^2 = 0.03$). The four independent variables contributed another 0.05 in shared variability. In total 14% of the variability in naming performance was predicted by scores on these four independent variables.

8.1.3.3 Discussion of predictor variables

In the two forms of analyses conducted here to attempt to identify the effect of a number of psycholinguistic variables on SC's word retrieval there was evidence from both forms of analysis of the influences of animacy and of frequency. SC is better at naming non-living than living entities, and better at high frequency items⁴¹. In the matched subsets analysis familiarity was significant, but the regression analysis failed to find an independent effect of this variable. The regression analysis revealed a further effect, that of age of acquisition, which the matched subsets failed to identify. This variable correlates highly with both frequency and familiarity. It is therefore difficult to ascertain the independent effect of each of these three variables. The two variables which had an effect in both forms of analysis were animacy and frequency. In summary it is safe to say that SC's naming is influenced by animacy, but also by lexical factors including frequency.

8.1.4 Written picture naming

8.1.4.1 Method and Results

A subset of the 200 picture items was selected for the assessment of written naming. SC was unable to attempt any of these, being unable to generate any letters at all. He was then asked to name the picture orally and then try to spell the word. This made no difference to his performance on written naming which was still at floor, regardless of whether he could name the item orally. He was able to copy letters and words, write part of the alphabet when asked to write this out without any dictation (17/26 letters produced), write some letters to dictation (8/10), but was unable to spell words to dictation.

8.1.4.2 Interpretation of results

The data from this assessment indicate that SC is unable to access the orthographic output lexicon from semantics (written picture naming) and from auditory input (spelling to dictation). He is also unable to access orthography from phonology as shown by his inability to spell words he can name. Whether representations within the

⁴¹ In the matched subsets analysis frequency was not examined in isolation from familiarity.

orthographic output lexicon remain intact but inaccessible, or whether these are damaged and unavailable remains an open question.

8.1.5 Phonological and orthographic cueing of picture naming

The procedure for cueing picture naming is outlined in Chapter Five, section 5.4. The results for SC are shown in Table 8.9.

TABLE 8.9 SC Cueing of picture naming

Condition	Phonological	Orthographic
Extra time (n = 12)	2	1
Single cue (n = 12)	4	0
Choice of cues (n = 12)	5	0

Chi square analyses of the phonological cueing, comparing the three conditions was not significant (chi square = 1.83, df = 2, n.s.). A second analysis comparing the choice condition with extra time was also not significant (chi square = 0.81, df = 1, n.s.). As there was clearly no effect of orthographic cues over extra time, no statistical analysis was conducted on the data.

8.1.6 Interpretation of naming data

SC's attempts at naming a set of 200 words on two occasions, and a related study investigating his ability to respond to phonological and orthographic cues, reveal the following main findings.

SC has a severe anomia with percentage scores of around 0.35 on repeated tests of picture naming. He makes mainly semantic errors or fails to name an item at all. The nature of the semantic errors, being mainly circumlocutions, suggest that SC has knowledge of the target semantics. The presence of associate errors and, to a lesser degree, other lexical substitutions, suggests SC may have a semantic deficit. SC's ability to name an item is affected by animacy with non-living things being more available than living things. This further supports the notion of a semantic deficit. In addition to a semantic deficit SC may also have damage to the phonological output lexicon: his naming performance was highly significantly predicted by frequency and to a lesser extent familiarity and age of acquisition.

SC failed to respond to phonological or orthographic cues at all. This indicates that there may be some damage to representations in the output lexicon: cues fail to lower the threshold (logogen account) or raise the activation level (Dell et al's 1997 account) of the target sufficiently to allow activation of the phonological form. This implies that there may be items SC will never name whose threshold is so high or whose resting level of activation is so low that no input stimulus can enable adequate activation to be achieved.

This has implications for therapy. The first phase of therapy involves the administration of phonological or orthographic cues. As SC is unable to respond to a single administration of such a cue it is unlikely that repeated administration over a number of sessions will have a different effect on his naming.

8.2 SEMANTIC PROCESSING

SC's semantic processing was tested on three tests of semantics: Pyramids and Palm Trees, and spoken and written word to picture matching.

8.2.1 Pyramids and Palm Trees

The three picture version of this test was used to investigate non-lexical semantic processing. SC scored 46/52 correct, which is outside the range of normal controls, but yet does not imply a severe semantic deficit.

8.2.2 Word to picture matching

Spoken and written word to picture matching tests were administered. SC scored 26/30 in the spoken version, and made four semantic errors. He completed the task easily and quickly. Normal controls make a maximum of two errors on this task and SC is therefore just outside the normal range of performance.

SC found written word to picture matching much more difficult. He appeared very uneasy in carrying out the task, and reported he could not make sense of the written

word⁴². He said on several occasions that he would “take a gamble” in making his choice, that his mind was telling him it’s wrong and that he did not like the choice he was making. He tried at times to name all the pictures in the array in order to eliminate those which did not match the written input. Nevertheless he managed to score 23/30. His errors were either semantic (4/7) such as selecting teapot for target kettle, or phonological (3/7) such as selecting hoof for roof. Given his overall score and his error pattern in this task, SC’s discomfort with the task indicates that he is processing more information than he is conscious of. The presence of semantic errors indicates that SC is accessing the correct semantic field in most cases. In three cases he selected the phonological distractor, although in all cases a visual processing deficit or a deficit in transcoding visual input to phonological output may have been the cause of the error, e.g. he read boy aloud as ‘ball’.

8.2.3 Interpretation of results

The evidence from the three tasks carried out indicates that SC has a deficit in semantics. His performance was outside the range for normal controls on all three tasks, and he made semantic errors in both word to picture matching tasks. SC is able to access semantics more easily via the auditory modality than via the written modality. SC’s processing of written stimuli suggest a deficit in processing at the level of visual analysis or at the level of the visual input lexicon, coupled with a semantic deficit. Given his unease and lack of confidence in his responses in this task he scored highly however, suggesting a processing routine which operates quite well but of which he is unaware, or for which he is monitoring operations and judging the performance too stringently.

8.3 AUDITORY DISCRIMINATION

8.3.1 Test performance on minimal pairs

Auditory discrimination was tested using the Action for Dysphasic Adults non-word minimal pairs. SC struggled to understand this task, and could not process the stimuli from the tape recording. The test was therefore administered by the researcher live. SC

⁴² The stimuli were enlarged for this assessment to ensure that SC’s visual condition did not contribute to the outcome.

scored 38/40 on this test. This was carried out without lip-reading. Of the errors both were false positives (calling a different pair same).

8.3.2 Interpretation of results

The results from this task indicate that SC has good auditory analysis skills under certain conditions. When there are too many variables he is unable to make judgements on incoming stimuli. When these variables are reduced - in this instance by eliminating the different voices - performance is good.

8.4 SHORT TERM MEMORY

8.4.1 Test performance

This was tested with four different tests (see section 5.3.3). SC scored as follows on these tests: digit span 2.5; letter span 2.2, phoneme span 2.3; picture pointing span 2.7.

8.4.2 Interpretation of results

SC is compromised in short-term auditory memory. The good performance in the modified test of auditory analysis implies that the apparent short term memory deficit is not due to input auditory processing problems. This deficit is apparent in all tests of this function to a comparable degree indicating that this relates to a general problem in storing and reproducing (apart from pointing span) auditory-phonological codes.

It is possible to relate this impairment to SC's poor response to phonological cues. SC showed no significant response to phonological cues which may be explained as an inability to maintain the cue in short term memory and use it to operate upon word finding. This would be particularly marked in the case of choice of phonological cues where two cues were involved. SC's phoneme span is 2.3 however and therefore may be sufficient to allow him to use cues. In which case the lack of response to cues may relate to an overall deficit in the phonological output lexicon (as described in section 8.1.6). This second interpretation is supported by the parallel data from orthographic cues where again no effect of cues was found. It is unclear how auditory short-term memory relates to use of cues.

In the related therapy study (for methods see Chapter Five, for results see Chapter Ten) participants were asked to respond to a choice of up to four phonological cues (and in the orthographic condition, four orthographic cues). For example for target ‘cake’ they would be told “It begins with buh, puh, kuh or luh”. It is probable that SC will also find this aspect of therapy difficult and this element may make the therapy ineffective for him.

8.5 READING ALOUD AND REPETITION

8.5.1 Reading aloud words

SC found this task extremely difficult. He was unable to complete the full set of 182 items, and eventually a subset were extracted in their place which were controlled for frequency imageability and length. In reading the larger set SC took a long time over each item, and talked about many of the words he could not read aloud saying that they

TABLE 8.10 SC’s distribution of errors in reading aloud words

Error type (n = 44)	Raw score	% score
Visual then semantic error	3	6.5
Visual error	7	15.2
Provided spoken definition	1	2.2
Gestured use of item	2	4.3
No response / comment on word	31	67.4

were words, or were ‘perfect’ or ‘alright’, but that he could not say anything. Most of his errors were of this type although he also made some semantic errors, some visual errors, some mixed visual and semantic errors, and in one case provided a definition of the word. In the subset he scored 8/52 correct. In this set he made no semantic errors.

TABLE 8.11 SC’s errors in reading aloud

Error type	Target	Response
Visual then semantic error	plea	happiness (via pleasure?)
Visual error	worm	woman
Provided spoken definition	monkey	a little animal. man had it on a shoulder

The distribution of SC’s error types is shown in Table 8.10. Examples of the various error types are shown in Table 8.11. Of the eight items which SC managed to read aloud

all were in the high imageability subset, most (6/8) were also high frequency, and all were one or two syllables in length with no three syllable words read aloud correctly.

8.5.2 Reading aloud non-words

SC found this task extremely difficult and expressed discomfort at having to undertake it. The task was abandoned after five items. SC's responses to these five items consisted of comments upon the visual stimulus such as saying an item was "not a word" or was "wrong" indicating that although he was unable to translate from orthography to phonology he was aware of the illegality of the stimulus.

8.5.3 Repetition of words

The set of 182 words were presented for SC to repeat. SC achieved 103/182 correct (57%). His errors were mainly phonologically related to the target. Details of the distribution of errors is given in Table 8.12. Responses were categorised as phonologically related or not. To qualify as phonologically related a response had to share 50% or more of its phonemes with the target.

The data demonstrate that most of SC's responses are correct (103/182) or are related phonologically to the target (49/182), indicating that he retains some ability to convert input phonemes to spoken output. Within the set of errors he is more likely to produce a non-word (65 errors, 82% of all errors) than a word (14 errors, 18% of all errors).

TABLE 8.12 SC repetition of words distribution of error types

Error type (n = 79)	Raw score	%
Phonological error: word	11	13.9
Phonological error: non-word	38	48.1
Unrelated word	3	4.0
Unrelated non-word	27	34.2

The subset of the 182 words controlled for frequency imageability and length (n = 52) were then analysed. Analysis of this subset revealed a significant effect of length (one syllable: 16/20; two syllables: 11/20; three syllables: 4/12: Jonckheere Trend Test, $z = 2.45$, one tailed $p = 0.007$) but no effect of frequency or imageability.

8.5.4 Repetition of non-words

SC scored a total of 7/26 correct when the criterion is accurate production of the whole phoneme string. As for word repetition there was a significant effect of length with SC getting six one-syllable stimuli, one two-syllable stimuli, and no three-syllable stimuli correct (Jonckheere Trend Test, $z = 2.56$, one tailed $p = 0.005$).

Closer analysis reveals better performance than this score would suggest. SC produced 13/26 correct initial phonemes, and when the total set of phonemes are considered he scored 86/130 (0.66) correct (correct phonemes in the correct order).

8.5.5 Interpretation of reading and repetition results

SC has a severe deficit in trans-coding orthography to phonology, demonstrated by his difficulties in completing the tests of reading words and non-words aloud. His complete inability to read non-words aloud indicates that the sub-lexical routine translating graphemes to phonemes is not operating. This means he is reliant upon the lexical route to read words aloud. The presence of visual and semantic errors, although these are not plentiful, support the notion of reading via a damaged lexical semantic route. This hypothesis is supported further by the fact that the few items he was able to read aloud were high in imageability and frequency, which are thought to operate at semantic and lexical levels respectively. Recall also that SC made semantic and visual/phonological errors in word to picture matching⁴³.

There is also some evidence of a length effect in reading aloud, although the data are extremely limited. Combined with the evidence from the short term memory tasks, this might support a claim for output buffer damage. As SC's reading aloud is so impaired the data are limited and therefore strong conclusions cannot be drawn.

SC also has a deficit in trans-coding input phonology to output phonology in repetition tasks. His good performance in auditory analysis implies that any deficit found here is not due to an input processing impairment. He is able to repeat some non-words

⁴³ Like KR whose data are presented in Chapter Nine SC has deep dyslexia. Further assessments to look at function word reading and derivational errors in reading would provide more information on this aspect. As SC's reading is so impaired and this aspect was not the focus of the present study further assessments were not undertaken.

indicating that there is at least some retained function in auditory-to-phonological conversion. Non-word repetition is affected by the length of the stimulus, providing evidence for an output buffer impairment. Word repetition is also impaired. His errors, which were mainly phonologically related to the target and mainly non-words, and the significant effect of word length on repetition, indicate that SC is relying to a large extent on the sub-lexical route to repeat words. Either a deficit in this route, or a deficit in the output buffer, means that repetition via this route is impaired. In addition the preponderance of non-words in his word repetition implies that self-monitoring systems are impaired⁴⁴. There is no evidence of variables reflecting lexical semantic involvement affecting SC's word repetition.

8.6 TESTS OF INTERNAL PHONOLOGY

8.6.1 Test performance

On the picture homophone test (where the person has to identify the two items which have the same phonology) SC reported he could not do this task and displayed some distress with his difficulties. The task was therefore abandoned after five items. Of the five he got one correct.

On a test of initial phoneme knowledge, where the person has to select the item which has the same initial phoneme as the target item, SC scored 5/20. For those which he got right he was reliant upon spoken output to guide his choice and was unable to complete the task silently.

8.6.2 Interpretation of results

These tests of internal phonology are notoriously difficult for people with aphasia. However, SC's performance is at floor on both tasks indicating that he either failed to grasp the task or that he has no access to internal phonological knowledge or no conscious awareness of that knowledge.

⁴⁴ It is unclear in Patterson and Shewell's model how self-monitoring operates. Levelt et al (1991) proposed a post lexical editing mechanism to account for the tendency in normals to produce words rather than non-words in spoken word errors. How this is incorporated into most of the current models is not specified. Dell et al's (1997) account uses interactive activation to explain this phenomenon in normals. Pathologically rapid decay in the activation of nodes in aphasic speakers is used to explain non word errors. This creates incoherence between adjacent levels of processing.

8.7 SENTENCE COMPREHENSION

8.7.1 Test performance

In spoken sentence comprehension SC scored 10/16 correct, making errors on reversible active sentences (e.g. The soldier hits the singer), reversible passive sentences (e.g. The policeman is painted by the dancer) and on embedded sentences (e.g. The shoe under the pencil is red)

In written sentence comprehension SC had a lot of difficulty processing the stimuli and worked hard to parse the sentence. He scored 5/16 where 4/16 is chance. As for his auditory comprehension he made errors on reversible active sentences, reversible passive sentences and on embedded sentences. Unlike in his auditory comprehension however he made lexical semantic errors in the comprehension of simple SV and SVO sentences. For example he selected 'The woman is standing' for the target 'The woman is walking'.

8.7.2 Interpretation of results

SC is able to process simple SV and SVO sentences via the auditory modality. This is in line with his functional ability in everyday speech. He shows agrammatic comprehension with reversible sentences and more complex sentences involving embedded clauses.

SC's difficulties in sentence comprehension were similar across the two modalities, with marked difficulties on reversible actives and passives, and on embedded sentences. His comprehension of written sentences is also affected by a deficit in lexical semantic processing, which impacts on his understanding of simpler sentences (SV and SVO) as well. These sentences were processed well via the auditory modality. This is in line with the results from word to picture matching where SC made more semantic errors via written than via spoken input and indicates better access to semantics from auditory than from written input, with a possible semantic access deficit in the visual processing route affecting access from the visual input lexicon to semantics.

8.8 NON VERBAL PROCESSING

8.8.1 Non verbal problem solving

SC's score on a subset of the Raven's Progressive Matrices was 11/12, showing that visual problem solving was virtually intact.

8.8.2 Visual processing

SC was able to bisect horizontal lines successfully indicating that he was able to compensate for his hemianopia. He scored 27/32 on the BORB (Birmingham Object Recognition Battery: Riddoch and Humphreys, 1987) object decision task A hard which is the mean score for normal control participants.

8.8.3 Interpretation of results

SC's scores on the Raven's shows that he can carry out non verbal problem solving tasks. His performance on the the BORB subtest shows that his processing of visual stimuli is intact. This was particularly important to establish given the presence of the degenerative disease affecting SC's vision.

8.9 CONVERSATION

The data considered here are from the two conversations produced at assessments one and two. These data also formed part of the analysis of the reliability of the conversation measure described in Chapter Six. In the discussion of SC's data, in particular the qualitative analyses, the issue of the reliability of subjective interpretation of the data needs to be borne in mind. Nevertheless it is possible to draw some tentative conclusions about SC's language processing difficulties and how these affect his conversation.

8.9.1 Turn taking

SC's wife LC was his conversation partner throughout the five assessments. In the two conversations SC produced both substantive and minimal turns, and the majority of his

turns were substantive. He is able to initiate topics and maintain topics both through statements and asking questions.

Conversation Sample 8.1

- 1 SC I know . it's beautiful . I met him many times . you've been once
- 2 LC have I
- 3 SC yes
- 4 LC you're sure [it's east
- 5 SC [of a summer
- 6 we just just walk in . it's so expensive
- 7 LC expensive . in Kent
- 8 SC no
- 9 LC Essex . that's all that's east=
- 10 SC =no . further
- 11 LC you can't go further east . you'll be in the North Sea (laughs)
- 12 SC no . or else the other way then (7.22)
- 13 LC so it is west
- 14 SC yes
- 15 LC it is west
- 16 SC yes
- 17 LC see I was thinking Cambridge [and Essex
- 18 SC [sorry

PA: person with aphasia

LC: conversation partner

Extract from SC's first conversation

In both of the conversations considered here there was a lot of turn overlap between the two speakers. SC's turns involved multiple attempts at an utterance construction and it was difficult for LC to project turn completion and subsequent transition relevance points. This led to LC overlapping with SC, as she believed he was reaching an end of a turn when in fact he was part way through. In assessment one the conversation sample consisted of SC trying to find the name of a place he wished to visit and LC trying to guess what this may be. The conversation therefore consisted of one long repair sequence. SC's attempts to help LC guess were often misleading as he made

unmonitored errors such as ‘yes’ for ‘no’, and semantic errors which went unchecked. LC produced a number of behaviours to manage SC’s aphasia such as prompting production of a target which she knew.

Conversation sample 8.1, which is from SC’s first conversation, shows both elements referred to in the outline above. In a previous turn SC has responded that the place he is thinking of is east of London. In fact he means west and this semantic error continues to throw the conversation off course for several turns. He continues to confuse east and west, even on hearing LC refer to west, until line 12, where he realises the error. This evidence of semantic errors in input and output is in line with the data from formal testing. Some of the difficulties in overlap are also evident here. In line 4 LC is trying to orient SC to the possibility that east may not be correct. He misses this however as he is still completing his previous turn from line 1 with a reference to the time of year that LC went to the place in question (“in the summer”). The combination of the production of semantic errors in SC’s speech, which are not checked, with a failure to process that same error when it is repeated back to him by LC, and the presence of overlaps which cause some of the conversation to be unperceived by SC, lead to frequent breakdowns in mutual understanding, which take multiple turns to repair.

8.9.2 Lexical retrieval

SC produces a range of items from different lexical classes in conversation. He produced nouns, verbs, adjectives adverbs and numerals. When word types are considered, that is the number of distinct individual lexical items he produces within any given grammatical class, the noun category scores highest. SC produces many verbs but these tend to be repeated use of high frequency verbs such as go, come, get, say.

8.9.3 Trouble indicating behaviours

SC produced semantic errors, phonological errors and neologisms in both conversations⁴⁵. As shown above, word finding difficulties formed a significant cause of the breakdowns in SC’s conversations and led to long repair sequences. Comments on word-finding difficulties were also common. SC did not tend to produce long pauses in his conversations, tending instead to carry on talking when stuck for a word. His ability

⁴⁵ Recall that the inter-rater reliability for identification of paraphasias was not strong (Chapter Six) and therefore the data are offered in this light.

to generate comments about word-finding and produce fluent automatic speech ensured that the conversation flow was maintained even though the business of that conversation was often the repair of a breakdown.

8.9.4 Repair

As shown above a number of factors relating to SC's language processing and his conversational behaviour contributed to the occurrence of breakdowns in understanding in conversation between him and his wife. On the whole these breakdowns were repaired by a collaborative and often long sequence between SC and his wife. LC made multiple guesses as to the intended target, and SC either rejected or accepted these. Although repair was successful on the taped samples, LC reported that this was not always the case in other conversations.

8.9.5 Conversation measure

The variables found to have good inter-rater reliability and good test-retest reliability in the conversation measure (see Chapter Six) made up the final version of the measure. This was then used to analyse SC's data at assessments one and two⁴⁶. These data are shown in Table 8.14.

SC produced more speech than LC in the two conversations although there is little difference in terms of the number of turns produced. SC's turns are thus longer than LC's and this is in part due to LC's facilitating SC's productions. Both speakers produce substantive and minimal turns, although the proportion of these differs across the two conversations with both speakers producing more minimal turns in the second conversation. Numbers of content words and of nouns are comparable across the conversations. SC's lexical retrieval difficulties are not apparent from these data as only filled pauses showed sufficient reliability to be included here.

⁴⁶ The data here are from the analyses carried out by the author of this report.

TABLE 8.14 Conversation measure: data for SC at assessments one and two

VARIABLE		Assessment One	Assessment Two
B	NUMBER OF SPEECH UNITS		
B1	Person with aphasia	401	478
B2	Conversational partner	323	201
C	Turn taking		
C1	Number of turns: Person with aphasia	66	51
C2	Number of turns: Conversational partner	66	46
C3	Number of substantive turns: Person with aphasia	49	32
C4	Number of substantive turns: Conversational partner	48	23
C5	Number of minimal turns: Person with aphasia	5	12
C6	Number of minimal turns: Conversational partner	5	19
D	Word retrieval and speech errors (PA only)		
D1	Total number of content words (excluding paraphasias)	101	125
D2	Total number of nouns (subset of content words)	38	33
D11	Number of filled pauses within PA's turn	6	3
E	Repair		
E1	Instances of repair	1	4
F	Proportional data		
F1	Substantive turns / turns (PA)	0.74	0.63
F2	Minimal turns / turns (PA)	0.08	0.26
F3	Content words / speech units	0.25	0.26
F4	Nouns / speech units	0.09	0.07
F5	Filled pauses / speech units	0.01	0.006
F6	Instances of repair/total turns	0.01	0.04

8.9.6 Discussion of SC's conversation data

Despite significant expressive language problems SC is a lively and creative contributor in conversation. The samples analysed here suffer from the 'observer's paradox' in that the speakers' behaviour is influenced by the presence of the recording device. This is evident from LC's attempts to promote and facilitate SC's contributions. The data show however that SC is able to initiate and maintain topics, to use substantive and minimal turns, to signal when misunderstandings occur, and to contribute to repair when breakdowns occur. Breakdowns are due to word-finding difficulties, a failure to repair misunderstandings, and sentence construction difficulties. SC produces nouns in conversation, and semantic errors, phonological errors and neologisms.

8.10 SUMMARY OF SC'S LANGUAGE PROFILE

SC presents with expressive and receptive aphasia. He is able to maintain a conversation, and has sufficient comprehension to allow him to participate in most everyday conversations. His expressive problems centre around word finding and sentence construction difficulties.

Word finding is severely impaired. SC scored at around 0.35 in the picture naming assessment. His errors were mainly semantic errors and failures to respond. The semantic errors were associates of the target or circumlocutions. The presence of semantic errors suggests a possible semantic deficit. This hypothesis is supported by the fact that the semantic variable animacy predicts SC's word finding. He is better at naming non-living than living items. A further effect of lexical variables, in particular word frequency, was found however, suggesting that word finding is also breaking down in the access to or within the phonological output lexicon. There is no data from written naming as SC was unable to attempt this at all.

Evidence from cueing of word retrieval supports the notion that representations in the output lexicon are damaged or are not available. SC's word finding was not affected by the administration of cues through either the spoken or the written modality. Further evidence to support the hypothesis of damage to the output lexicon comes from his impaired word repetition (0.57) and severely impaired word reading (0.15).

The hypothesis that SC has a semantic deficit was supported by the results from tests of input semantics. Access to semantics from picture materials was impaired. There was a discrepancy between written and spoken word to picture matching, with spoken word to picture matching showing a mild deficit, and written word to picture matching being more impaired.

Further evidence for a semantic deficit comes from SC's reading aloud of words. SC cannot read non-words at all, although he is aware that they are illegal. Thus he is reading via the lexical semantic route. His reading aloud is severely compromised, with 0.15 correct. He makes semantic and visual errors, and there is a possible imageability effect. A reasonable conclusion from these data is that SC is reading via a damaged lexical semantic route, with possible semantic involvement.

The functioning of SC's auditory input processing and auditory to phonological output was assessed. SC is able to judge non-word minimal pairs successfully, indicating that his auditory analysis system is operating well. He can repeat short non-words well although struggles with longer words. This partially preserved ability and the presence of a length effect indicate that the auditory-phonological conversion mechanism is functioning to a large extent but is limited as to the number of phonemes it can manage. This system is therefore available to support word repetition. SC can repeat some words although this function is impaired. His error pattern in word repetition differs from that found in naming and reading aloud. He makes no semantic errors, produces mainly phonological approximations to the target, and has a marked length effect. This finding supports the contention that he is repeating words via the sub-lexical conversion system⁴⁷. The finding of a length effect suggests that there may be damage to the output buffer, although this is not supported by findings from either picture naming or reading aloud. A second possibility is that the length effect emerges as a result of an overall incapacity on the part of the conversion mechanism.

The findings from the assessments carried out above have implications for the type of therapy that may be effective for SC. It is unlikely that the first phase of therapy which involves cueing of target words, will aid SC's naming. On the other hand, the second phase of therapy, in which there is much more opportunity for participants to retrieve words of their choice in appropriate semantic contexts may be more beneficial.

⁴⁷ This contention begs the question of why there is a difference in SC's processing of words vs. non-words. This difference could emerge because of a difference in the difficulty of stimuli, or there could be some residual activation of the lexical semantic route providing support to words, or there could be some activation of output phonology,

CHAPTER NINE: ASSESSMENT RESULTS FOR KR

9.0 INTRODUCTION

In this chapter assessment results for KR are presented. Some personal background information is provided, followed by an analysis of tests of language processing, including word finding in picture naming and in conversation. The chapter follows the format adopted for the previous two chapters. For further details of the design and the methods used please refer to Chapter Five.

9.0.1 Background details

At the time of entering the study in 2000 KR was 39 years old. She had sustained a single left hemisphere CVA 13 years previously. KR was born in Liberia and spoke Liberian and English fluently prior to onset. She was educated up to the age of 20 and prior to the onset of her aphasia she worked as a secretary. As a result of the stroke KR has aphasia, and a right quadriplegia affecting the upper limb. She walks unaided but is unable to use her right arm and hand for everyday activities, and writes with her non-preferred left hand. She no longer works but manages to bring up her two children independently. Throughout the course of her participation in this study she attended group speech and language therapy with a focus on functional adaptation and communication. She attends keep-fit classes and swims regularly. KR lives in her own home with her husband and two children. Her husband acted as her conversation partner throughout the study.

KR: Cinderella narrative

*the brothers um sisters . two brothers you uh mix . tidy . everything . and um . one day
the . the . the cinder . one day he leaves . er you understand . he want to go . and um the
fairy . he want to go fairy . and um /paŋkɪn / and um two . um rats*

KR has a non-fluent aphasia. Her connected speech sample (above) demonstrates a short phrase length, errors and omissions in syntactic structure, and many false starts and revisions. Her word finding difficulty is apparent in connected speech and conversation, where lexical errors, for example 'brothers' for target sisters, 'rats' for target mice, and phonological errors for example /paŋkɪn / for pumpkin, and omissions

occur. Nevertheless KR is an effective communicator, using speech, gesture and facial expression to convey her thoughts.

9.1 PICTURE NAMING

9.1.1 Scores at baseline naming

The set of 200 black and white line drawings described in previous chapters was used to assess picture naming. For details of this set see Chapter Five (section 5.2). KR's scores at assessments one and two are shown in Table 9.1.

TABLE 9.1 KR's scores in picture naming at assessments one and two

Assessment One		Assessment Two	
Raw score (n=200)	%	Raw score (n=200)	%
80	40	74	37

9.1.2 Error analysis

Details of the responses KR made in picture naming are shown in Table 9.2.

TABLE 9.2 KR's Picture naming responses in assessments one and two

Response	Assessment One		Assessment Two	
	Raw score	%	Raw score	%
Correct	80	40	74	37
Visual error	7	4	5	3
Semantic error	19	9	21	10
Phonological error	2	1	2	1
Non-words	3	2	2	1
Rejections	0	0	7	4
No response	89	44	89	44
Other	0	0	0	0
TOTAL	200	100	200	100

The distribution of correct responses and the various errors types is consistent across the two assessment times. The majority of KR's errors were failures to respond. She also made a number of semantic errors. Other error types were sporadic only. Examples of KR's errors are given in Table 9.3.

TABLE 9.3 Examples of KR's naming errors

Error type	Target	Response
Visual error	river	boat*
Semantic error	tank	war
Phonological error	tweezers	/pizə/
Non-words	tractor	chock
Rejections	grass	tree no

*All of KR's visual errors involved naming a small item in the picture rather than something visually related to the target, thus do not classify as visual perceptual errors.

Examples of the different semantic error types are shown in Table 9.4.

TABLE 9.4. Examples of KR's semantic errors

Error type	Target	Response
Superordinate	tulip	flowers
Semantic co-ordinate	triangle	square
Semantic associate	buckle	belt
Circumlocution	clown	funny people
Semantically and visually related	pie	cake
Semantic then phonological	muzzle	jog (via dog)
Semantically and phonologically related	button	buckle

TABLE 9.5 Distribution of KR's semantic error types in assessments one and two

Semantic error type	Assessment one		Assessment two	
	Raw score	%	Raw score	%
Superordinate	0	0	4	19
Semantic co-ordinate	5	26	3	14
Semantic associate	8	43	3	14
Circumlocution	1	5	0	0
Semantically and visually related	5	26	9	43
Semantic then phonological	0	0	1	5
Semantically + phonologically related	0	0	1	5
Total semantic errors	19	100	21	100

The majority of KR's semantic errors were semantic associates of the target, and semantic and visually related errors. The distribution of the error types is shown in Table 9.5.

9.1.3 Psycholinguistic variables influencing KR's naming

As for the previous two participants, two analyses were conducted here. The set of 200 pictures contained matched subsets for a number of variables whose influence could thus be analysed independently: imageability, animacy, operativity, familiarity, familiarity and frequency combined, age of acquisition, and length (number of syllables). In each analysis a number of items had been assigned to the two matched subsets (ranging from 30 per set to 55 per set), thus the majority of the data were excluded from each individual analysis. This form of analysis is therefore weakened by the number of exclusions, but does allow the analysis of the influence of one variable in isolation from other variables

The second analysis used multiple regression to identify the contribution of the variables to performance. Ellis, Lum, and Lambon-Ralph (1996) and Greenhouse, Bromberg and Fromme (1995) describe and provide a critique of and examples of applications of these techniques to single case data.

9.1.3.1 Matched subsets

The results of the analyses for the matched subsets are shown in Table 9.6. The data shown are the result of analysing the two assessment times combined. The latter measure was used as this involved more data and was therefore more powerful than consideration of one assessment time alone.

TABLE 9.6 KR analysis of psycholinguistic variables in matched subsets

Variable	A1 plus A2	
	z score	p value
Imageability	0.31	0.38
Animacy	1.09	0.14
Operativity	1.19	0.19
Familiarity	2.01	0.02*
Familiarity and frequency	2.95	0.002**
Age of acquisition	-1.77	0.04*
Length	0.44	0.33

Table 9.6 shows z-scores and values of p for one-tailed tests. * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$

The raw data for all variables showing a significant effect in the analysis are shown in Table 9.7. This table shows numbers correct at assessments one and two only.

TABLE 9.7 KR numbers correct in matched subsets

Variable		n =	Assessment one		Assessment two	
			Score	%	Score	%
Familiarity	High	30	16	53	14	47
	Low	30	7	23	9	30
Familiarity and frequency	High	50	28	56	22	44
	Low	50	11	22	13	26
Age of acquisition	Early	40	18	45	18	45
	Late	40	11	28	10	25

This analysis shows that there is an effect of familiarity, with high familiarity items being named more readily than low familiarity items. There is a highly significant effect of frequency and familiarity combined. Items which are of high familiarity and frequency are named more easily than those which are low on these variables. Finally there is an effect of age of acquisition.

These analyses indicate that a set of coherent variables predict KR's naming success. All three variables identified here are probably lexical in origin⁴⁸. There is no effect from any of the semantic variables imageability and animacy, but there is a tendency towards an effect of operativity at assessment one, although this does not reach significance. There is no effect of length indicating that there is no impairment of post lexical processes in KR's naming.

This analysis did not consider the following variables: concreteness (which correlates highly with imageability) and frequency (which correlates highly with familiarity). In KR's case, where all three variables found to have an effect upon word finding were lexical in origin it is likely that frequency will also affect naming. In the following analysis, regression techniques were used to identify the influence of the variables included in the above analyses plus concreteness and frequency.

⁴⁸ For details of the probable loci of the effects of the different variables see Chapter One, section 1.4.

9.1.3.2 Regression analysis

The set analysed here is the same set as that used to analyse PH and SC's naming in Chapters Seven and Eight respectively. An inter-correlation matrix was computed to identify co-variance in the variables involved, and the relationship between each variable and the dependent variable. The dependent variable is naming success measured by the number of times an item was named correctly by KR in two administrations of the naming test, and thus ranges from 0 to 2. This inter-correlation matrix is shown in Table 9.8.

The variables which correlate significantly with KR's naming are: operativity, familiarity, frequency and age of acquisition. In the case of familiarity, frequency and age of acquisition this finding mirrors that shown by the analysis of matched subsets. The impact of operativity was not found in the matched subsets analysis. Thus one semantic variable and three lexical variables appear to predict KR's naming. Familiarity appears to have the strongest impact. Although there are significant correlations between the independent variables, most notably between familiarity, frequency and age of acquisition, it was decided to include all three in the regression. The correlations, although high, are not extremely so.

**TABLE 9.8: Inter-correlation matrix for predictor variables
for the naming set for KR**

	IMAG	CONC	OPER	ANIM	FAM	FREQ	AOA	SYLL	PHON
IMAG	1.00	-.360**	.073	.101	.141	.034	-.131	.213**	.231**
CONC		1.00	.356**	.137	-.158*	.007	.274**	.034	.059
OPER			1.00	.278**	.325**	-.047	-.201**	-.013	-.065
ANIM				1.00	.197**	.113	.152*	-.063	-.041
FAM					1.00	.672**	-.509**	-.114	-.178
FREQ						1.00	-.406**	-.081	-.195**
AOA							1.00	.418**	.448**
SYLLS								1.00	.843**
PHON									1.00
KR NAMING	.129	-.073	.241**	.068	.415**	.283**	-.288**	-.067	-.111

Values are for Pearson's R. IMAG = imageability, CONC = concreteness, OPER = operativity, ANIM = animacy, AoA = age of acquisition, FAM = familiarity, FREQ = Celex log combined frequency, SYLLS = number of syllables, PHON = number of phonemes. Significance levels of R are ** $p \leq 0.01$, * $p \leq 0.05$.

A simultaneous regression analysis was carried out with naming performance as the dependent variable and operativity, familiarity, frequency and age of acquisition as the independent variables. R for regression was significantly different from zero ($F = 10.957$, $df = 4, 183$, $p = 0.000$) and overall R^2 was 0.193. One of the independent variables contributed significantly to the prediction of naming performance: familiarity ($t = 2.664$, $p = 0.008$; $sr^2 = 0.03$). One further variable approached significance: operativity ($t = 1.745$, $p = 0.083$; $sr^2 = 0.01$). Together the four independent variables contributed a further 0.15 in shared variability. In total 19% of the variability in naming performance was predicted by scores on these four independent variables.

9.1.3.3 Discussion of predictor variables

The two analyses both indicated that familiarity, frequency⁴⁹ and age of acquisition predicted KR's naming. The correlation matrix suggested additionally that frequency in isolation predicted naming success, along with operativity. As there are strong relationships between these four variables it is difficult to draw out the independent contribution of each. The regression analysis including all four variables showed an effect only of familiarity suggesting that this is the main predictor variable. Lesser effects of the other three were present but with familiarity included in the analysis no independent effects were significant. Taking evidence from both sources it is safe to conclude that KR's naming is strongly affected by lexical variables and most notably by familiarity, and that there is a possible effect of the semantic variable operativity.

9.1.4 Written picture naming

9.1.4.1 Method and Results

A subset of 40 of the 200 picture items was selected for the assessment of written naming. KR was shown the picture and asked to write down the name. Errors were coded as orthographic errors, unrelated errors, semantic errors and no responses. Orthographic errors were responses which contained at least 50% of the target's orthography in the correct order. Unrelated errors were responses which did not match this criterion for orthographic relatedness and could be words or non-words. Semantic errors were responses which did not match the orthographic criterion but had a semantic relationship with the target. Examples of the error types are shown in Table 9.9.

⁴⁹ In the matched subsets frequency was not analysed in isolation from familiarity however.

TABLE 9.9 KR written picture naming examples of errors

Error type	Target	Response
Orthographic: word	worm	worn
Orthographic: non-word	doctor	dotor
Unrelated: word	basket	but
Unrelated: non-word	kennel	pi
Semantic error	blouse	sleeve

The distribution of KR's error types is shown in Table 9.10.

TABLE 9.10 KR's written naming error distribution

Response	Number	%
Correct	12	30
Orthographic: word	4	10
Orthographic: non-word	13	32
Unrelated: word	1	3
Unrelated: non-word	2	5
Semantic error	3	8
No response	5	12
TOTAL:	40	100

KR produced mainly correct responses, or attempts at the target which maintained a significant percentage of the target orthography ($\leq 50\%$) but resulted in non-words (e.g. doctor \rightarrow dotor). These two response types (correct responses and orthographically related) combined show that in the majority of cases (72% of cases) KR accesses most of or all of the target orthography. This is better than her spoken naming performance. Of interest also is the fact that in 33/40 cases (82%) KR accessed the first letter of the target word.

9.1.4.2 Interpretation of results

The results outlined above imply that KR is able to access representations in the orthographic output lexicon but unable to produce the total set of letters in output, either due to impaired representations in the orthographic lexicon, or due to an impairment in maintaining the output from the lexicon in the graphemic buffer. Whatever the cause of the problems in producing written output, this assessment reveals that KR has access to a significant amount of orthographic information, which is in excess of the amount of phonological information available to her in the equivalent spoken form of the test.

9.1.5 Phonological and orthographic cueing of picture naming

The procedure for cueing picture naming is outlined in Chapter Five, section 5.4. The results for KR are shown in Table 9.11.

TABLE 9.11 KR Cueing of picture naming

Condition	Phonological	Orthographic
Extra time (n = 12)	2	3
Single cue (n = 12)	2	1
Choice of cues (n = 12)	2	2

The results of the phonological and orthographic cueing assessment show no effect of either the single cue or the choice of cues, when compared to the control condition of extra time. This is the case for both cueing conditions. The data indicate that, when unable to spontaneously produce a word, KR is not helped by either hearing the first sound or by seeing the first letter of the word. This further suggests that therapy which provides the same form of cueing, but in repeated attempts at the targets over a series of sessions, will also be ineffective. This aspect will be investigated in Chapter Ten.

9.1.6 Interpretation of naming data

KR's attempts at naming a set of 200 words on two occasions, and a related study investigating her ability to respond to phonological and orthographic cues, reveal the following main findings.

KR has a severe anomia with percentage scores of 0.37 and 0.40 on the two tests of picture naming. Her errors consisted mainly of failures to respond. Her spoken errors were mainly semantic, and were mainly single words related in meaning to the target. KR's spoken naming is affected by a number of lexical psycholinguistic variables. The data from the matched subsets indicate that familiarity, familiarity and frequency combined, and age of acquisition have a significant impact on KR's ability to retrieve a word.

This finding was partially corroborated by the regression analysis of the larger set of 188 words. Familiarity was found to significantly affect word finding. A lesser effect of age of acquisition, and a weak effect of frequency were also found. Unlike the findings

from the matched subsets the regression analysis also showed an effect for operativity. Taking the two sets of results together a parsimonious interpretation of the data would be that the main variable affecting KR's word finding is familiarity. The probable source of her word finding deficit is in the word level or in accessing that level (e.g. Dell et al, 1997: see Chapter Two, section 2.6.3.1).

KR also has a deficit in written naming but her error pattern differs from that seen in spoken naming: here she is able to access a significant amount of orthographic information and makes very few lexical errors. In an assessment of the effect of phonological and orthographic cues neither were more effective than extra time. In the next section KR's processing of semantic information is investigated further.

9.2 SEMANTIC PROCESSING

In order to investigate semantic processing KR was tested on three tests of semantics: Pyramids and Palm Trees (three picture version), and spoken and written word to picture matching.

9.2.1 Pyramids and Palm Trees

The three picture version of this test was used to investigate non-lexical semantic processing. KR scored 40/52 (0.77) correct (where normals' mean score is 0.99 and they make three errors or less). KR's score is thus far below the normal range of performance and suggest that she has a semantic processing deficit.

9.2.2 Word to picture matching

The CAT spoken and written word to picture matching tests were administered. KR scored 27/30 (0.90) in the written version, and 28/30 (0.93) in the spoken version. Normals make two errors at most. KR's errors involved selection of the semantically related foil (arrow for dart in both written and spoken tests, pencil for pen, tulip for rose, teapot for kettle).

9.2.3 Interpretation of results

The Pyramids and Palm Trees test is pictorial and therefore not a test of lexical semantics. In addition it relies significantly upon culturally-specific encyclopaedic knowledge of the world (e.g. that pigs and not donkeys eat acorns, that apples and not onions grow on trees). As KR was brought up in Liberia her semantic encyclopaedic knowledge may differ considerably from that of a European native, particularly with relation to crops and wildlife. It is important to bear this in mind when interpreting the test results and, although her test score is outside normal limits, further corroboration is required before declaring a semantic deficit to be present.

KR's performance is within normal limits for spoken word to picture matching and just outside normal limits for written word to picture matching. Her errors were semantic in nature and suggest at most a mild semantic impairment. However, the semantic foils she selected are also visually similar to their targets, suggesting a possible visual-perceptual impairment. Against this hypothesis is the fact that the visual errors she produced in picture naming involved naming a sub-part of the picture rather than a visually related item. There is some further evidence against this proposal from written naming, where she made no visual errors. This aspect was investigated further through tests of visual perception (see section 9.8.2).

9.3 AUDITORY DISCRIMINATION

9.3.1 Test performance on minimal pairs

Auditory discrimination was tested using the Action for Dysphasic Adults non-word minimal pairs. KR scored 26/40 (0.65) which is significantly better than chance (Binomial Test, $p = 0.04$).

9.3.2 Error analysis

All her errors were false positives (saying two items were the same when they were different). Thus she scored 20/20 for the 'same' set and 6/20 for the 'different' set.

9.3.3 Interpretation of results

In carrying out the test of minimal pair discrimination KR perceived very few differences between the 'different' stimuli pairs, indicating an impairment in phoneme discrimination. This deficit does not affect access to meaning from auditory input however: in spoken word to picture matching KR never chose the phonologically related foil and scored within normal limits on this test. It is possible that the auditory input stimulus combined with the pictorial information from the array of pictures enabled KR to make the correct choice in word to picture matching. An impairment in auditory discrimination may explain the lack of an effect of phonological cues on picture naming however.

9.4 SHORT TERM MEMORY

9.4.1 Test performance

This was tested with four different tests (see section 5.3.3). KR scored as follows on these tests: digit span 2.5, letter span 1.9, phoneme span 1.7, picture pointing span 2.1 .

9.4.2 Interpretation of results

Like most people with aphasia KR has impaired auditory short-term memory. She did better on the two tests involving real words: numbers and pictures. She fared worse on the phoneme and letter span tests, where no semantic information supports processing.

Her phoneme span score of 1.7 suggests a possible cause of her poor performance on the auditory discrimination task. In the latter task the person has to retain the two stimuli for long enough to make a judgement about their similarity. With such a short span for this form of input it is perhaps not surprising that KR did so badly on the auditory discrimination task. Thus one could hypothesise that her ability to analyse phonemes in words is intact (as she does not make phonological errors in word to picture matching), but that her short-term store for non-word phonological information is impaired.

KR's lack of response to phonological cues can also be explained in line with this hypothesis. In addition to a difficulty in identifying the phoneme or in discriminating

two phonemes (in the case of choice of cues), KR may fail to retain the information long enough in her short-term store for effective access to the appropriate phonological form to be achieved.

9.5 READING ALOUD AND REPETITION

9.5.1 Reading aloud words

KR scored 119/182 correct (0.65). Her errors were mainly visual and failures to respond, and some semantic errors. The distribution of her error types is shown in Table 9.12.

TABLE 9.12 Distribution of KR's response types in reading aloud

Response	Number	%
Correct	119	0.65
Visual error	21	0.12
Semantic error	5	0.03
Semantic plus visual error	3	0.02
Morphological error	4	0.02
Non-words	7	0.04
Other (unrelated word, perseveration, rejection of correct)	5	0.03
No response	18	0.09
TOTAL	182	1.00

When totalled together KR's semantic and semantic plus visual errors give a total of 8 responses or 5% of the total.

A subset of 52 of the 182 words made up matched subsets controlled for frequency, imageability and length. KR showed no effect of frequency (19/36 high, 16/36 low: chi square = 0.22, df = 1, n.s.), or of length (one syllable 14/20, two syllables 12/20, three syllables 8/12: Jonckheere Trend Test, $z = 0.14$, one tailed $p = 0.44$) but she showed a marked effect of imageability (30/36 high, 5/16 low: chi square = 11.39, df = 1, $p = 0.001$).

9.5.2 Reading aloud non-words

KR scored a total of 0/26 correct when the criterion is accurate production of the whole phoneme string. Using any other criterion (first phoneme correct, number of phonemes

correct) she fared as badly as she was unable to access phonology from orthography at all.

9.5.3 Repetition of words

KR scored 163/182 correct (0.90). Her errors were mainly phonological (where $\geq 50\%$ of the target phonemes are retained in the correct order), and non-word responses.

9.5.4 Repetition of non-words

KR scored a total of 18/26 (0.69) correct when the criterion is accurate production of the whole phoneme string. She showed a tendency to a length effect although this was not significant (9/10 one syllable, 6/10 two syllable and 3/6 three syllable: Jonckheere Trend Test, $z = 1.51$, $p = 0.07$ n.s.).

Closer analysis revealed better performance than the overall score would suggest. KR produced 21/26 (0.81) correct initial phonemes, and when the total set of phonemes are considered she scored 114/130 (0.88) correct phonemes in the correct order.

9.5.5 Interpretation of reading and repetition results

The semantic errors in KR's reading aloud suggest she has deep dyslexia. Coltheart (1987) stated that deep dyslexia is defined by the presence of semantic errors in reading aloud, and if a person makes semantic errors in reading single words they will also produce the following symptoms in reading aloud: visual errors, function word substitutions, derivational and inflectional errors, an inability to read pronounceable non-words, an imageability effect, function words read aloud worse than content words, impaired writing and spelling. KR produces semantic errors (although the number is not large), she makes visual errors, derivational errors, and shows a deficit in spelling. She cannot read aloud non-words. Analysis of a subset of the 182 words revealed an effect of imageability on reading aloud with high imageability words read aloud better than low⁵⁰.

⁵⁰ This is unlike KR's picture naming where mainly lexical variables affect word retrieval. The reason for this disparity is not clear. Presumably in both tasks semantic mediation is required. One would therefore predict an effect of imageability in picture naming as well as in reading aloud. One explanation lies in the restricted range of imageability in picture items.

This set of symptoms provides strong evidence for the presence of deep dyslexia⁵¹.

In contrast to her reading aloud KR showed good repetition skills. Repetition of words was mildly impaired, with the majority of errors being phonologically related to the target. This is unlike her picture naming where very few phonological errors were produced, and suggests reliance on direct correspondences between the input and output lexicons. Repetition of non-words was impaired but KR showed some retained ability to translate from input phonology to output phonology without semantic or lexical mediation. This is in contrast to her poor performance on auditory discrimination and her extremely reduced phoneme span.

9.6 TESTS OF INTERNAL PHONOLOGY

9.6.1 Test performance

The picture homophone test asks the person to judge silently the phonology of three picture items and decide which two have the same phonology (e.g. flower and flour). KR was completely unable to carry out this task silently, and when attempting it with overt naming managed only one item and was still unable to make the judgement independently.

On a test of initial phoneme knowledge, where the person has to select the item which has the same initial phoneme as the target item, KR was again unable to carry out the task.

9.6.2 Interpretation of results

KR's inability to carry out both tasks suggests a lack of access to meta-linguistic knowledge, or a failure to understand the tasks. Her good repetition of words indicates however that phonological output lexicon representations are intact and available.

⁵¹ No analysis of KR's function word reading was conducted as this was not the focus of this investigation. Further tests would involve function word reading aloud, and investigation of derivational and inflectional errors.

9.7 SENTENCE COMPREHENSION

9.7.1 Test performance

In spoken sentence comprehension KR scored 10/16 correct. She was able to interpret simple subject-verb sentences, non-reversible subject verb object sentences and subject verb and prepositional phrase sentences. Of the five reversible sentences in the test she scored one correct. She scored two out of four on embedded sentences. In written sentence comprehension she scored 9/16 correct. The error pattern was similar to that seen in spoken sentence comprehension with difficulties in reversible sentences and embedded sentences.

9.7.2 Interpretation of results

KR shows a difficulty in extracting meaning from word order as shown by her problems with reversible as opposed to non-reversible sentences, and her difficulty with embedded sentences. This agrammatic sentence comprehension mirrors her agrammatic connected speech production.

9.8 NON VERBAL PROCESSING

9.8.1 Non verbal problem solving

KR's score on a subset of the Raven's Progressive Matrices was 6/12, showing that visual problem solving was impaired.

9.8.2 Visual processing

KR was able to bisect horizontal lines successfully indicating that there was no hemianopia. She scored 26/32 on the BORB (Birmingham Object Recognition Battery: Riddoch and Humphreys, 1989) object decision A hard task where the normal range is 22-30 correct. She correctly accepted 15/16 real objects and correctly rejected 11/16 unreal objects.

9.8.3 Interpretation of results

KR's scores on the Raven's Progressive Matrices indicate a difficulty with visual problem solving. Her score on the object decision subtest of the BORB was within the normal range and indicates that she does not have an impairment in visual processing. There are conflicting results regarding this aspect of KR's processing. In line with a hypothesis of a visual perceptual impairment is the fact that she selected visually related distractors in word to picture matching tasks. However there is no evidence from spoken or written naming to support there being such an impairment.

9.9 CONVERSATION

The data considered here are the two conversations produced at assessments one and two. These data also formed part of the analysis of the reliability of the conversation measure described in Chapter Six. In the discussion of KR's data, in particular the qualitative analyses, the issue of the reliability of subjective interpretation of the data needs to be borne in mind. Nevertheless it is possible to draw some tentative conclusions about KR's language processing difficulties from the conversation samples and to speculate about how these affect her conversation.

9.9.1 Turn taking

KR produces equal numbers of turns as her partner and her turns are both substantive and minimal, with substantive turns dominating. She uses both statements and questions to initiate topics and topics shifts and to contribute to topic maintenance. It is important to note however that the topic under discussion determines to a large degree the extent to which KR can contribute to the conversation. When the topic is centred around events in her and her family's lives, or is initiated by her (e.g. in conversation two she recounts the details of a friend's trip to the US in which she initiates many topic shifts) KR is able to contribute significantly and meaningfully to the conversation. When the topic is initiated by her partner and concerns an abstract issue such as politics KR is unable to participate fully.

In the following sample KR is telling her conversation partner about Pancake Day at the children's school.

Conversation Sample 9.1

- 1 KR and uhm I'm so happy
- 2 SR [mmm
- 3 KR [you know I got fun
- 4 SR ok you happy because you helped to prepare the [pancake
- 5 KR [prepare the pancake day
- 6 SR ok
- 7 KR and uhm (3 seconds pause) uhm . it was fun it was fun it was fun
- 8 SR how you do it . the children was there you=
- 9 KR =no no no uhm uhm just uhm uhm just adults adults
- 10 SR oh I=
- 11 KR =two adults two adults
- 12 SR mmm
- 13 KR me and uhm just [adults] two person two adults
- 14 SR [mmhm]
- 15 KR it was fun I enjoy it
- 16 SR that's good

KR: person with aphasia

SR: conversation partner

Extract from KR's first conversation

In the sample KR produces substantive turns to introduce topic elements and to respond to her partner's queries effectively.

9.9.2 Lexical retrieval

In sample 9.1 KR produces a number of nouns spontaneously (fun, adults, person, although note the plural error in line 13 'two person'). She also accesses verbs in conversation and some adjectives and adverbs.

9.9.3 Trouble indicating behaviours

KR produces some semantic errors, and some phonological and neologistic errors in conversation. In the following sample her partner is explaining that after a day's work he is feeling hungry. KR encourages him to go and make breakfast when the target is

dinner. Examples such as this where a clear semantic error is produced by KR are relatively rare however. The main way in which her difficulties in word retrieval become apparent is through the large number of filled pauses produced in her turns (see sample 9.1 line 9 above). There are also a number of lengthy unfilled pauses in her turns although filled pauses preponderate. It is not always clear what the source of the problem is when a filled pause is produced by KR. It is possible that sentence construction difficulties are also a cause of these breakdowns.

Sample 9.2

- 1 SR so I mean I'm really hungry now I spend all day working all the time
- 2 KR oh
- 3 SR I wanted something of =
- 4 KR = alright alright alright go and make breakfast

KR: person with aphasia

SR: conversation partner

Extract from KR's first conversation

9.9.4 Repair

The large number of filled pauses leads to frequent interjections by KR's partner in an effort to solve the apparent problem, or attempts by him to find out more about what KR is trying to say. The latter type is the most common form of repair in these conversations. There are frequent overlaps in turn-taking, as collaborative repair sequences emerge involving both KR and her partner.

Sample 9.3

- 1 KR yeah yeah you know the the the uhm he said uhm he's a Muslim
- 2 SR oh
- 3 KR er Muslim and uhm he's from uhm =
- 4 SR = Kashmir

KR: person with aphasia

SR: conversation partner

Extract from KR's second conversation

KR is unable to complete her turn in line 3 and the difficulty (which may be to do with memory, knowledge or a word finding difficulty) is signalled to her partner by the filled pause. He interjects with the word he believes she is searching for in his next turn.

In the next sample the partner seeks clarification twice as to KR's movements that day. In this sample KR 'shadows' his speech as she attempts to corroborate his statements. This leads to overlap in turn-taking.

Sample 9.4

- 1 SR so lunchtime you came home =
- 2 KR = home yes yeah
- 3 SR you came [home] lunchtime =
- 4 KR [həʊ]
- 5 KR = time yeah

KR: person with aphasia

SR: conversation partner

Extract from KR's first conversation

9.9.5 Conversation measure

The variables found to have good inter-rater reliability and good test-retest reliability in the conversation measure (see Chapter Six) made up the final version. This was then used to analyse KR's data at assessments one and two⁵². These data are shown in Table 9.13.

KR and her partner produce equal numbers of turns, and although the partner produces more speech units KR's totals are still fairly high. In conversation one KR produces many more substantive than minimal turns. In conversation two the discrepancy between the two turn types is less marked. This may be due to the topic under discussion in conversation two: KR's partner introduced a political topic to which KR contributed merely minimal turns. The proportion of speech units that are content words is comparable across the two assessment times, although the proportion of nouns is

⁵² The data here are from the analyses carried out by the author of this report.

lower in the second conversation. Filled pauses are copious and are stable across the two times.

TABLE 9.13 Conversation measure: data for KR at assessments one and two

VARIABLE		Assessment One	Assessment Two
B	NUMBER OF SPEECH UNITS		
B1	Person with aphasia	246	299
B2	Conversational partner	340	350
C	Turn taking		
C1	Number of turns: Person with aphasia	69	54
C2	Number of turns: Conversational partner	68	55
C3	Number of substantive turns: Person with aphasia	46	32
C4	Number of substantive turns: Conversational partner	46	42
C5	Number of minimal turns: Person with aphasia	13	20
C6	Number of minimal turns: Conversational partner	22	11
D	Word retrieval and speech errors (PA only)		
D1	Total number of content words (excluding paraphasias)	76	98
D2	Total number of nouns (subset of content words)	46	32
D11	Number of filled pauses within PA's turn	37	37
E	Repair		
E1	Instances of repair	8	4
E	Proportional data		
F1	Substantive turns / turns (PA)	0.67	0.59
F2	Minimal turns / turns (PA)	0.19	0.37
F3	Content words / speech units	0.31	0.33
F4	Nouns / speech units	0.19	0.11
F5	Filled pauses / speech units	0.15	0.15
F6	Instances of repair / turns	0.06	0.04

9.10 SUMMARY OF KR'S LANGUAGE PROFILE

KR presents with an agrammatic aphasia with marked word-finding difficulties. The agrammatism is evident in production, where she has difficulties in verb retrieval and sentence construction, and in comprehension where she shows a marked difficulty with reversible sentences.

KR has a word finding difficulty. This is evident across a range of spoken word production tasks including picture naming, narrative production and conversation. Although there may be a mild semantic deficit, which is suggested by the fact that her performance on Pyramids and Palm Trees is outside normal limits, and that one semantic variable operativity predicted naming, it is more likely that the focus of her word finding problem is the retrieval of the phonological form.

KR's word finding is greatly affected by lexical variables, in particular familiarity. This provides evidence for a deficit in accessing phonological representations. Moreover in written picture naming the vast majority of her responses are either correct or orthographically related to the target indicating that she can access orthographic representations from (intact) semantics. The integrity of KR's phonological representations is attested to by her good performance on word repetition (this despite poor minimal pair discrimination). It is likely that a mild semantic deficit coupled with a phonological access problem underlie KR's word retrieval problems.

Attempts to cue KR's spoken word production were ineffectual. There are a number of explanations for this. She scored at chance on a test of non-word minimal pair discrimination indicating a deficit in the auditory analysis system. She showed an impaired auditory short term memory span, and she showed a deficit in repetition of non-words. The mechanism by which cues operate is poorly understood (see section 4.2.1.2) but must involve auditory analysis of phonemes. It may be that some activation of the phonological representation must occur for a cue to have an effect, and that for KR resting activation was too low for the cue to work. In this explanation KR may well respond to cues using an easier set of words. Alternatively, evidence from assessment of auditory short term memory identified a severely impaired phoneme span. In order to use a cue effectively the person has to be trying to retrieve the phonology as well as processing the auditory input. In the test carried out here both single cues and a choice of cues were tried. Both were equally ineffective. If the problem were due to impaired short term memory one might expect a single cue to work better than a choice of cues. KR's deficit in repetition of non-words was not as marked as expected given her inability to judge non-word minimal pairs, and her poor phoneme span. This residual ability may mean that there is some justification in attempting to treat words through phonemic cues in the subsequent therapy experiment.

KR's word finding problems are apparent in all tasks involving spoken word production. She is able to produce some nouns in all tasks but makes some semantic errors, and often fails to produce a target word at all. In picture naming this is marked as a failure to respond and in narrative production and conversation filled pauses are evident throughout. KR's attempts to produce the Cinderella story showed that she has access to nouns in connected speech. Despite her marked aphasia KR can initiate and maintain topics in conversation. She produces both substantive and minimal turns, although the nature of the topic under discussion may affect the proportion of these two turn types. She can access nouns in conversation, but has frequent difficulties with this. When she fails to retrieve a noun she usually produces a filled pause. This signals trouble to her partner and repair strategies are then initiated.

CHAPTER TEN: THERAPY RESULTS

10.1 DESIGN AND METHODS OF THERAPY STUDY

10.1.1 Participants

The three participants described in the previous chapters all entered both phases of therapy. The results of the language assessment and investigation of cueing were used to predict each individual's response to the therapy administered in each phase. For ease of reference the results of key assessments for all three participants are compiled in Table 10.1.

TABLE 10.1. Assessment results for PH, SC and KR

Task	n =	Participant		
		PH	SC	KR
Picture naming assessments 1 and 2 (mean)	200	0.36	0.32	0.40
Semantic tests:				
CAT Spoken word to picture matching	30	0.93	0.87	0.93
CAT written word to picture matching	30	0.97	0.77	0.90
Pyramids and Palm Trees three pictures	52	0.90	0.88	0.77
Proportion semantic errors	-	0.25	0.28	0.16
Phonological tests:				
ADA auditory discrimination test	40	0.68	0.95	0.65
Short term memory test (phoneme span)	-	2.50	2.30	1.70
Repetition of words	152	0.97	0.57	0.90
Repetition of non-words	26	0.58	0.27	0.69
Repetition of non words: initial phoneme correct	26	0.88	0.50	0.81
Proportion phonological errors	-	0.05	0.02	0.02
Reading aloud words	152	0.97	0.15	0.64
Reading aloud non-words	26	0.35	0.00	0.00
Reading non words initial phoneme correct	26	0.85	0.00	0.00

The table shows participants' performance on the following: CAT Comprehensive Aphasia Test (Swinburn et al., in preparation); Pyramids and Palm Trees (Howard and Patterson, 1992); ADA Auditory Discrimination from Action for Dysphasic Adults Comprehension Battery (Franklin et al., 1992). The remaining assessments are unpublished. Proportion of errors are the total of the error type divided by the total number of errors.

10.1.1.1 Participant PH

PH presents with a severe anomia in the context of otherwise relatively preserved single word processing. She also shows a deficit in decoding sentence meaning from syntactic information. Her word finding deficit appears to arise from a semantic processing

deficit which is relatively mild, and from a lexical access deficit. The semantic deficit is apparent from three lines of evidence. PH shows a deficit in one test of input processing (Pyramids and Palm Trees), she produces some co-ordinate errors in spoken output in picture naming, connected speech, and in conversation (although see caveats regarding the reliability of the conversation data), and her picture naming is predicted by semantic variables (concreteness and animacy, and possibly operativity). Evidence for this being a mild deficit comes from the fact that her performance in verbal input tasks is within normal limits, and that she mainly produces semantic information about the target in spoken picture naming, with co-ordinate errors being more rare. This indicates that she can access semantic information but fails to access phonology.

PH's phonological representations appear to be intact and available as shown by her good performance in transcoding from input to spoken output when semantic mediation is not required in reading aloud and repetition, and from the evidence of successful word retrieval when a phonological or orthographic cue is provided. Support for the hypothesis that PH has a difficulty in mapping from semantics to phonological output representations comes from the fact that her naming performance is associated with age of acquisition. It is therefore proposed that PH has a mild semantic deficit and a further deficit in mapping from semantic representations onto phonological forms.

PH responded well to both phonological and orthographic forms of cueing and was able to benefit from the choice of cues as well. As the therapy administered in Phase One uses the same method as that used in this cueing investigation it is likely that this form of therapy will be effective for PH. In the second phase of therapy participants were encouraged to retrieve target words in more naturally occurring speech, with support from pictorial material and cues as required. PH is able to retrieve a certain number of nouns in conversation, to signal when she is unable to find a word, and to benefit from cues, thus again it is likely that she will benefit from this form of therapy.

10.1.1.2 Participant SC

SC presents with expressive and receptive aphasia. He is able to maintain a conversation, and has sufficient comprehension to allow him to participate in most everyday conversations. His expressive problems centre around word finding and sentence construction difficulties.

Word finding is severely impaired. SC scored around 0.35 in the picture naming assessment. The presence of semantic errors in naming suggests a possible semantic deficit. This hypothesis is supported by the fact that the semantic variable animacy predicts SC's word finding. He is better at naming non-living than living items. The possibility that SC has a semantic deficit was supported by the results from tests of input to semantics. Further evidence for a semantic deficit comes from SC's reading aloud of words where he makes semantic and visual errors, and has a possible imageability effect.

A strong effect of lexical variables, in particular word frequency, but also familiarity and age of acquisition, was also found in picture naming, suggesting that word finding is breaking down in access to or within the phonological output lexicon. Evidence from cueing of word retrieval supports the notion that representations in the output lexicon are damaged or are not available. SC's word finding was not facilitated by the administration of cues through either the spoken or the written modality. Further evidence to support the hypothesis of damage to the output lexicon comes from his impaired word repetition and severely impaired word reading.

Given the possible dual impairments in semantics and within the phonological output lexicon, and the lack of a cueing effect, it seems unlikely that the first phase of therapy which involves cueing of target words, will aid SC's naming. SC can find some words in conversation however, and is able to use the language he has creatively to work his way around problems with word-finding when these arise. The second phase of therapy, in which there is much more opportunity for participants to retrieve words of their choice in appropriate semantic contexts, may be more beneficial for him.

10.1.1.3 Participant KR

KR presents with an agrammatic aphasia with marked word-finding difficulties. The agrammatism is evident in production, where she has difficulties in verb retrieval and sentence construction, and in comprehension where she shows a marked difficulty with reversible sentences.

KR has a word finding difficulty. Like PH, her word finding problem appears to arise from two sources: she has a mild semantic deficit and a lexical access problem.

Evidence from a number of sources indicate that she has a deficit at the semantic level. She has a deficit in accessing semantics from pictures, and scores just outside normal limits on written word to picture matching. She makes semantic errors in picture naming and in narrative production and conversation, and her word finding is possibly affected by operativity which is a semantic variable. She also makes semantic errors in reading aloud.

KR's word finding is greatly affected by lexical variables in particular familiarity. This suggests a deficit in accessing phonological representations. The integrity of KR's phonological representations is attested to by her good performance on word repetition (this despite poor minimal pair discrimination). However, attempts to cue KR's spoken word production were ineffectual.

It is likely therefore that a mild semantic deficit coupled with a deficit in lexical retrieval underlie KR's word retrieval problems. This is similar to PH who has a mild semantic deficit, but differs from SC who has a more marked semantic deficit and in addition has damage to the phonological output lexicon. It is unclear therefore whether repeated exposure to cues should help KR's word finding. The lack of any effect in the cueing experiment suggests the therapy will not work. On the other hand, as the phonological representations are accessible, for example in repetition, it is possible that repeated exposure to cues will activate representations sufficiently to facilitate access.

KR is able to participate in conversation, and can produce some nouns. Her conversation is punctuated by frequent word finding difficulties, shown by the large number of filled pauses and abandonment of turns. It is difficult to predict how she will respond to the open choice offered in Phase Two therapy, where participants are encouraged to select words freely in conversation with support from pictures and cues.

10.1.1.4 Summary

Table 10.2 summarises the main assessment findings with regard to the three participants, in terms of their semantic processing, the status of the phonological output lexicon, and their response to phonological and orthographic cues.

TABLE 10.2 Summary of the three participants

	PH	SC	KR
Semantic processing impairment	mild	moderate	mild
Phonological output lexicon	intact	impaired	intact
Response to cues	good	none	none

10.1.2 Method

The design of the therapy study is described in detail in Chapter Five, but brief details are repeated here for ease of reference. The study consisted of three phases each lasting approximately eight weeks. The first phase involved a period of assessment, and a facilitation experiment. The second and third phases involved two forms of therapy.

10.1.2.1 Therapy Phase One

The baseline assessment and facilitation phase was followed by the first phase of therapy (Phase One), which focused on improving word-finding in a picture naming task. Participants were seen once per week for a total of eight weeks, each session lasting roughly one to two hours. The effects of phonological and orthographic cues were evaluated in this phase (see Hickin et al, 2002a).

10.1.2.2 Therapy Phase Two

After this participants were invited to enter the second phase of therapy (Phase Two) which aimed to enable the person with aphasia to use treated words in tasks approximating closely to real-life conversation (reported in Herbert et al, 2003). After this phase of therapy participants were not seen for two months then were reassessed to provide follow-up data.

10.1.2.3 Assessments One to Five

Assessments one to five included a set of tests, repeated at key points throughout the study. There were two assessment points prior to therapy which acted as the baseline measure, one assessment point immediately after Phase One which tracked the effect of that phase of therapy, one assessment point after Phase Two which tracked the effects of that therapy, and a final assessment point after a period of no intervention, which

monitored the maintenance of therapy. Details of the set of tasks involved at each of the assessment points are given in Chapter Five and comprised picture naming of 200 items, a tape-recorded conversation, and a set of language control tasks.

10.2 EFFECTS OF THERAPY ON PICTURE NAMING

In analysing the naming data multiple comparisons using McNemar's chi square were carried out. Given the large number of statistical comparisons performed a conservative level of significance of $p = 0.01$ was used throughout. Full details of all test results are reported but only with levels of p at 0.01 or lower are treated as significant.

10.2.1 Effects of therapy on naming overall

10.2.1.1 Participant PH

The results for PH's picture naming at the five assessment points are shown in Figure

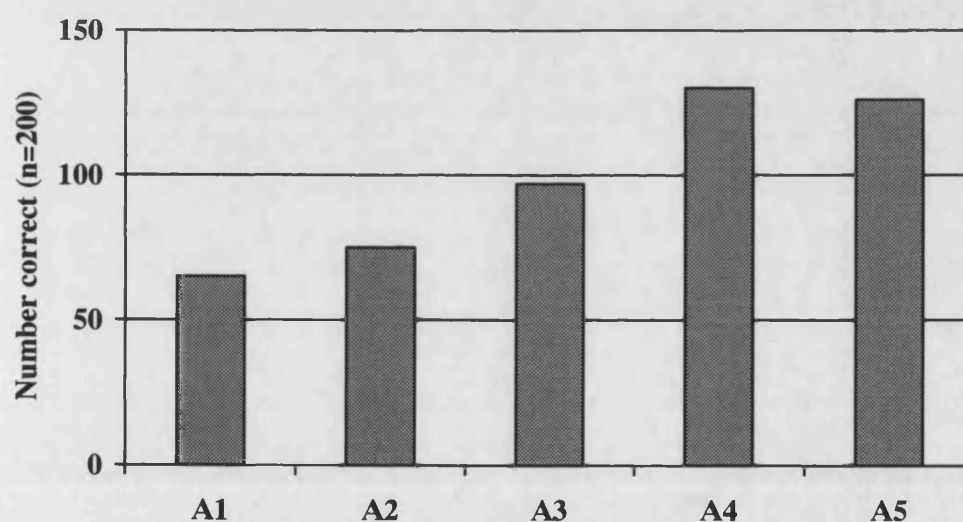


FIGURE 10.1 PH's Naming at Assessments One to Five

10.1. Statistical analyses were carried out to determine whether there was a significant change in performance between assessments one and two, two and three, three and four and four and five. The analyses showed no difference between overall scores at assessments one and two (McNemar, one tailed, $p = 0.34$, n.s.), a significant difference between assessments two and three (McNemar, one tailed, $p < 0.01$), a significant difference between assessments three and four (McNemar, one tailed, $p < 0.001$), and

no difference between assessments four and five (McNemar, one-tailed, $p = 0.34$, n.s.). There was therefore no change during the baseline testing phase and the maintenance phase, but significant overall change after each phase of therapy.

10.2.1.2 Participant SC

The results for SC's picture naming at the five assessment points are shown in Figure 10.2.

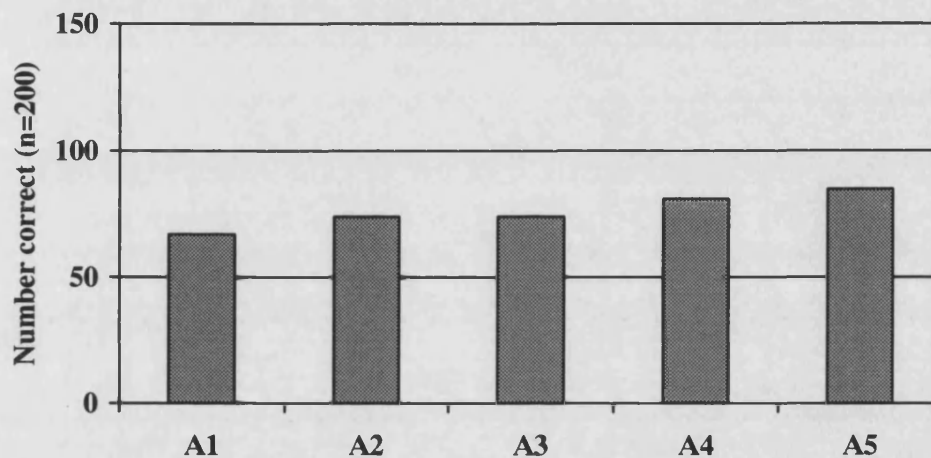


FIGURE 10.2 SC's Naming at Assessments One to Five

Statistical analyses were carried out to determine whether there was a significant change in performance between assessments one and two, two and three, three and four and four and five. McNemar analyses showed no difference between sets at any point (assessment one to assessment two: $p = 0.22$, n.s.; assessment two to assessment three: $p = 0.56$, n.s.; assessment three to assessment four, $p = 0.19$, n.s.; assessment four to assessment five, $p = 0.34$, n.s.).

10.2.1.3 Participant KR

The results for KR's picture naming at the five assessment points are shown in Figure 10.3.

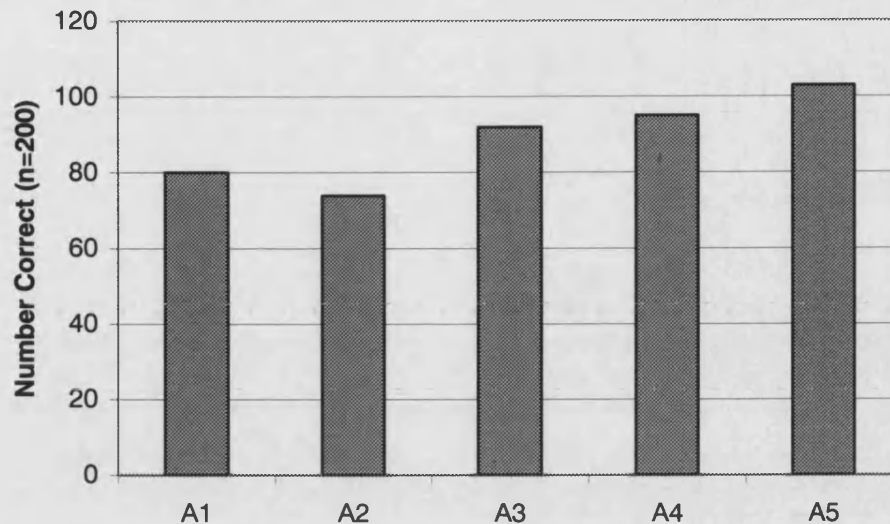


FIGURE 10.3 KR's Naming at Assessments One to Five

McNemar tests showed no difference between assessments one and two ($p = 0.23$), a significant difference between assessments two and three ($p = 0.01$), and no difference between assessments three and four ($p = 0.39$) and between assessments four and five ($p = 0.19$).

10.2.2 Therapy Phase One

10.2.2.1 Participant PH

10.2.2.1.1 Comparison of treated versus untreated sets

The 200 picture items were split into a treatment and a no treatment group. This allowed for the identification of item-specific effects and any generalisation of therapy effects onto untreated items. PH's performance in each of these two sets is shown in Figure 10.4.

The data indicate that the effect of the therapy was restricted to those items which underwent therapy, with no carryover of the treatment effect to the untreated set. McNemar tests showed that there was a significant difference for the treated set and no difference for the untreated set (treated set: $p < 0.001$; untreated set: $p = 0.26$, n.s.).

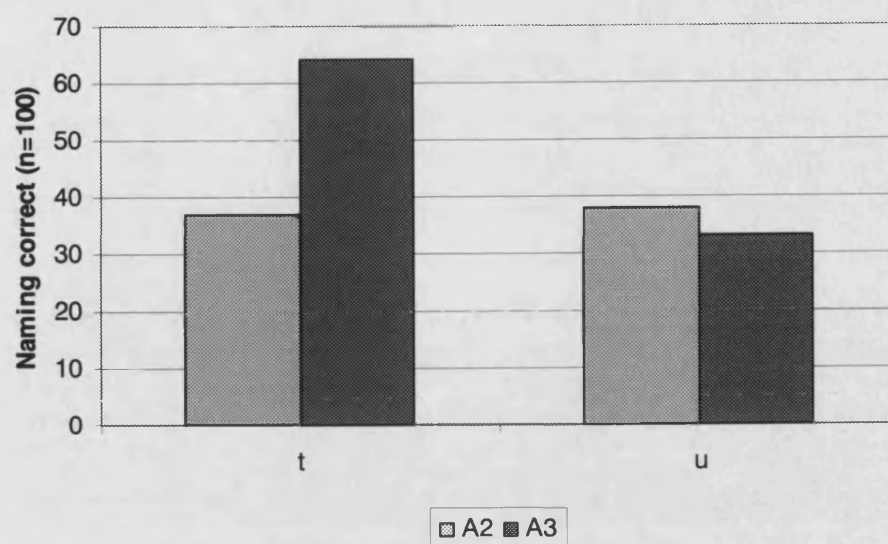


FIGURE 10.4 PH's naming of treated and untreated sets (Phase One) at assessments two and three

10.2.2.1.2 Comparison of phonological and orthographic approaches

The treated items were split into two sets for Phase One. One set underwent phonological and one set underwent orthographic therapy. Scores for the two treated sets and the untreated set are shown in Table 10.3.

TABLE 10.3 PH's scores for phonological and orthographic sets

	Phase One therapy			
	A2		A3	
	Score	%	Score	%
Phonological (n = 50)	18	0.36	30	0.60
Orthographic (n = 50)	19	0.38	34	0.68
Untreated (n = 100)	38	0.38	33	0.33

The orthographic set gained more items than the phonological set (15 versus 12). Both the phonological set (McNemar's test, $p = 0.0021$) and the orthographic set showed a significant change (McNemar's test, $p = 0.0007$).

10.2.2.1.3 Discussion of PH's Phase One results

The results from the analysis of the overall scores, which include all treated and untreated items, suggest that the baseline scores are stable thus there is no evidence of spontaneous recovery of picture naming, nor of any practice effects. The significant difference between scores at assessments two and three indicates an effect of therapy Phase One on the set as a whole. Likewise the difference between assessments three and four indicates an effect of Phase Two therapy. The lack of a difference between assessments four and five indicates that the therapy effect accumulated after both phases maintains at least up to two months post therapy. For PH there is a stable baseline, both phases of therapy were effective, and the therapy effect maintained. This is the most positive outcome one could envisage. The subsequent analyses were initiated to identify more discretely the nature of the therapy effects.

The analysis of treated and untreated sets for Phase One showed a therapy effect restricted to items undergoing therapy: there was no generalisation to untreated items. A comparison of gains made in the phonological and orthographic sets showed no difference between sets: both types of cueing therapy were equally effective for PH.

PH's main level of deficit lies between semantics and phonology, and she responded well to cues. According to Wheeldon et al (1992) the mechanism of the repetition priming effect in normal speakers lies in the links between semantics and phonology. The cueing and the therapy conducted here might reasonably be hypothesised to operate by strengthening those links. In PH's case therefore it is reasonable to assume that the cueing and the therapy operate in a similar fashion.

10.2.2.2 Participant SC

10.2.2.2.1 Comparison of treated versus untreated sets

One possibility with regard to SC's naming is that a therapy effect shown in the treatment set was offset by a loss of items from the no treatment set. In this case no overall change would be evident and a possible treatment effect would have been masked. This possibility was investigated by comparing his scores in the treated and untreated sets. SC's performance in each of these two sets is shown in Figure 10.5.

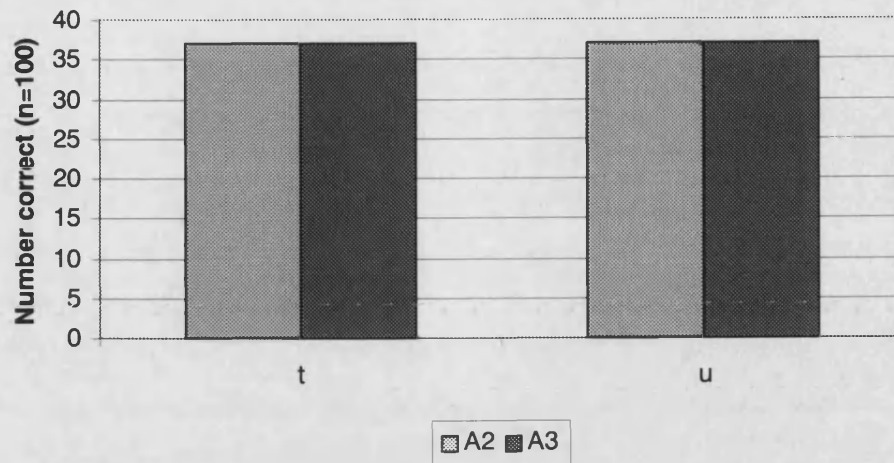


FIGURE 10.5 SC's naming of treated and untreated sets (Phase One) at assessments two and three

The chart shows that there was no effect of therapy at all for SC. Treated and untreated sets maintained their baseline scores.

10.2.2.2.2 Comparison of phonological and orthographic approaches

The treated items were split into two sets for Phase One. One set underwent phonological and one set underwent orthographic therapy. A further possibility is that within the set of treated items one set made gains whilst the other made losses. This possibility was assessed by comparing scores for the two treated sets and the untreated set. These are shown in Table 10.4.

TABLE 10.4 SC's scores for phonological and orthographic sets Phase One therapy

	A2		A3	
	Score	%	Score	%
Phonological (n = 50)	18	0.36	18	0.36
Orthographic (n = 50)	19	0.38	19	0.38
Untreated (n = 100)	37	0.37	37	0.37

The figures show that there was no independent effect on either of the two treatment sets.

10.2.2.2.3 Discussion of SC's Phase One results

The analysis of the overall set of 200 items showed no significant change across any of the assessment times. It can therefore be argued that the baseline was stable, that there was no overall effect of either therapy, and no change during the period of no intervention. Analysis of the treated and untreated sets showed that there was no item-specific effect of the therapy, and further analysis of the two treatment sets showed that neither the phonological nor the orthographic approach had an independent effect upon SC's naming.

The results support the prediction made regarding SC's naming and the effect of this form of therapy. SC showed no effect of cues when these were administered once and the repeated administration of the cues was similarly ineffective (see Chapter Eight, section 8.1.5).

10.2.2.3 Participant KR

10.2.2.3.1 Comparison of treated versus untreated sets

The 200 picture items were split into a treatment and a no treatment group. This allowed for the identification of item-specific effects and any carry-over of therapy effects onto untreated items. KR's performance in each of these two sets is shown in Figure 10.6.

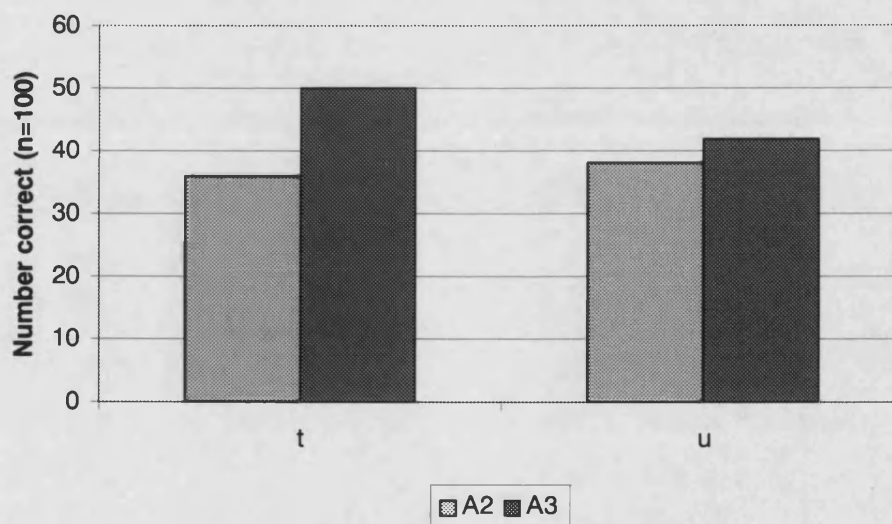


FIGURE 10.6 KR's naming of treated and untreated sets (Phase One) at assessments two and three

The chart indicates that the treatment effect was largely confined to treated items although there was some improvement in the untreated set. McNemar analysis of the change in the two sets showed a significant change for the treated set ($p = 0.008$), but no difference for the untreated set ($p = 0.27$, n.s.)

10.2.2.3.2 Comparison of phonological and orthographic approaches

The possibility of one type of cue being more effective than the other was analysed by comparing KR's performance in the two separate subsets and the untreated set. These data are shown in Table 10.5.

The figures show more improvement in the phonological than in the orthographic set. Further McNemar analyses showed a significant change in the phonological set only ($p = 0.007$) and a non-significant change in the orthographic set ($p = 0.23$, n.s.).

TABLE 10.5 KR's scores for phonological and orthographic sets
Phase One therapy

	A2		A3	
	Score	%	Score	%
Phonological (n = 50)	18	0.36	28	0.56
Orthographic (n = 50)	18	0.36	22	0.44
Untreated (n = 100)	38	0.38	42	0.42

10.2.2.3.3 Discussion of KR's Phase One results

The analysis of KR's naming of the total set of 200 words showed a stable baseline, an overall effect of Phase One therapy, no overall effect of Phase Two therapy, and no change during a period of no intervention.

Analysis of the treated and untreated sets showed a significant change in items treated in Phase One. Although there was a small improvement in the untreated set this was not significant. Thus, as for PH, the therapy affected treated items.

Investigation of the independent effects of the two types of cues showed that only the phonological approach led to significant gains for KR. The orthographic approach led to some improvement but this was not significant.

Like SC KR showed no effect of the administration of a single cue on her picture naming (see section 9.1.5). It is interesting therefore that, unlike SC, KR responded to at least the phonological cues. KR's language profile is similar to that of PH however, who responded well to cues and to Phase One therapy. KR's main deficit lies in accessing phonology from semantics. It may be the case that the therapy mechanism was the same as that which worked for PH – mappings between semantics and phonology were strengthened by repeated exposure to the targets and to the cues. A single administration of the cues was not however effective for KR. This may relate to underlying lower activation levels. In addition the therapy effect for KR was less marked than for PH, which indicates that more input is required for KR in order for the same effect to be achieved. This result demonstrates that for some people with aphasia a single assessment of the effectiveness of cues will not be sufficient to indicate whether a related therapy will be effective.

10.2.3 Therapy Phase Two

10.2.3.1 Participant PH

When the change in scores between assessments three and four was analysed for the total set of 200 items a significant difference between assessments three and four was found (McNemar, one tailed, $p < 0.001$) (section 10.2.1.1). This indicates that the Phase Two therapy had a significant effect on naming overall.

10.2.3.1.1 Comparison of treated and untreated sets

PH showed an item-specific response to therapy in Phase One. A comparison of therapy effects for treated versus untreated items for Phase Two (see allocation of items to sets in Chapter Five) showed the therapy effect was again restricted to those items undergoing treatment (McNemar test comparing assessments three and four: treated set: $p = 0.0000$; untreated set $p = 0.29$, n.s.). PH's performance in the two sets at assessments three four and five is shown in Figure 10.7.

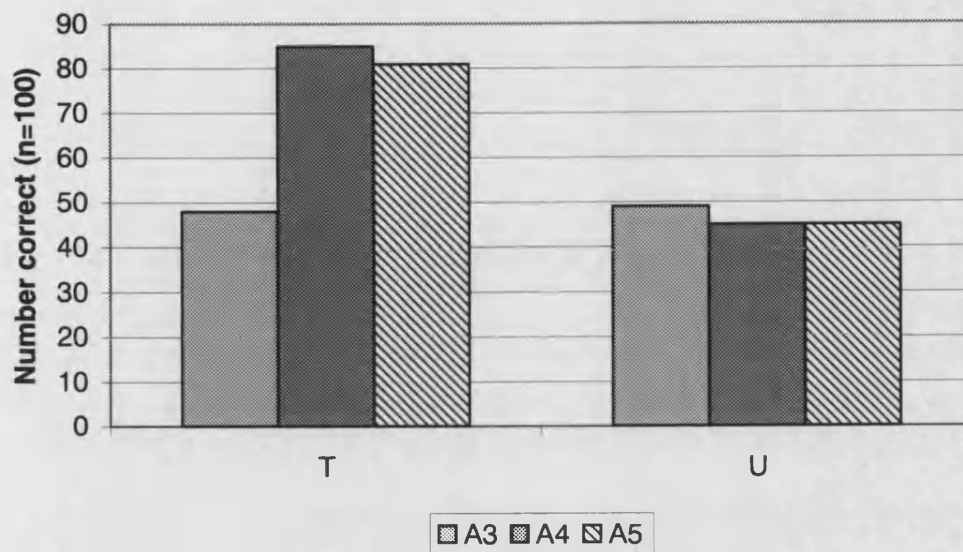


FIGURE 10.7 PH therapy Phase Two treated and untreated sets

10.2.3.1.2 Comparison of therapy effect across four subsets

Within the set of treated and untreated sets there are further subsets. Half of those items treated in Phase Two had already undergone treatment in Phase One: these items are termed Treated Treated or tT; half had been left untreated in Phase One, and are termed untreated treated or uT. Similarly for the items left untreated in Phase Two. Half had been treated in Phase One and are termed treated untreated or tU, and half had no treatment in either phase and are termed untreated untreated or uU.

Analysis of change in these subsets was carried out to determine where the therapy effect found in Phase Two was located. PH's proportion correct in the four subsets at assessments one to five are given in Table 10.6, along with significance levels for McNemar tests comparing scores at assessments three and four.

The set which was treated in both phases of therapy (tT) improved significantly after Phase Two therapy. The set which was treated in Phase One and left untreated in Phase Two (tU) showed a deterioration after Phase Two. The set which was treated in Phase Two only (uT) showed a significant improvement after Phase Two. The set which was untreated throughout showed no change after Phase Two.

TABLE 10.6 PH scores at assessments one to five for the four subsets of items

	A1	A2	A3	A4	A5	Value of p A3 to A4
tT(n=50)	0.32	0.36	0.62	0.84	0.78	0.004**
tU(n=50)	0.32	0.38	0.66	0.52	0.46	0.05*
uT(n=50)	0.32	0.38	0.34	0.86	0.84	0.000***
uU(n=50)	0.34	0.38	0.32	0.38	0.44	0.303

A1 to A5 = assessments one to five; tT = treated treated set, tU = treated untreated set, uT = untreated treated set, uU = untreated untreated set. Significance levels are: * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$. Significance levels are for McNemar tests comparing scores at assessments three and four.

The magnitude of the change in the set which was treated in Phase Two only deserves some consideration. The effect of Phase Two therapy only on this set of items was as great as the combined effect of both phases on the items treated in both phases (tT set).

There was no significant change in any of the four sets from assessment four to assessment five (McNemar tests comparing scores in each subset at assessments four and five: treated Treated: $p = 0.25$, n.s.; treated Untreated: $p = 0.27$, n.s.; untreated Treated: $p = 0.5$, n.s.; untreated Untreated: $p = 0.31$, n.s.).

10.2.3.1.3 Summary of PH's response to therapy Phase Two

There was an overall change in the 200 items as a result of Phase Two therapy. When the untreated and treated sets were compared the improvement was found to be restricted to treated items. Of the items treated in Phase Two both items treated in Phase one and those left untreated in Phase One showed a significant improvement as a result of Phase Two therapy. Items treated in both phases made as much change as those treated in just Phase Two. All four sets maintained at follow-up.

The mechanism proposed to be operating in Phase One - the mappings between semantics and phonology - may also be operating in this therapy for PH. Unlike in Phase One however, there was no explicit attempt to target processing at this level. Rather in Phase Two participants were encouraged to use the target words in interaction. This does not prohibit someone from using sound or letter knowledge in attempting to retrieve a word.

10.2.3.2 Participant SC

When the change in scores between assessments three and four was analysed for the total set of 200 items no difference was found between assessments three and four (McNemar, one tailed, $p = 0.19$, n.s.) (see section 10.2.1.2). This indicates that Phase Two therapy did not affect naming of the 200 items overall.

10.2.3.2.1 Comparison of treated and untreated sets

SC showed no response to therapy in Phase One. A comparison of therapy effects for treated versus untreated items for Phase Two showed a nearly significant effect for items treated in Phase Two (McNemar test: $p = 0.0113$; level of significance taken as $p \leq 0.01$), and no change in those items not treated (McNemar test: $p = 0.2024$, n.s.). SC's performance in the two sets is shown in Figure 10.8. The graph indicates that although the treated set T improved at assessment four, this may have been at the expense of the untreated set U which deteriorated at assessment four. The issue of maintenance investigated at assessment five will be dealt with below.

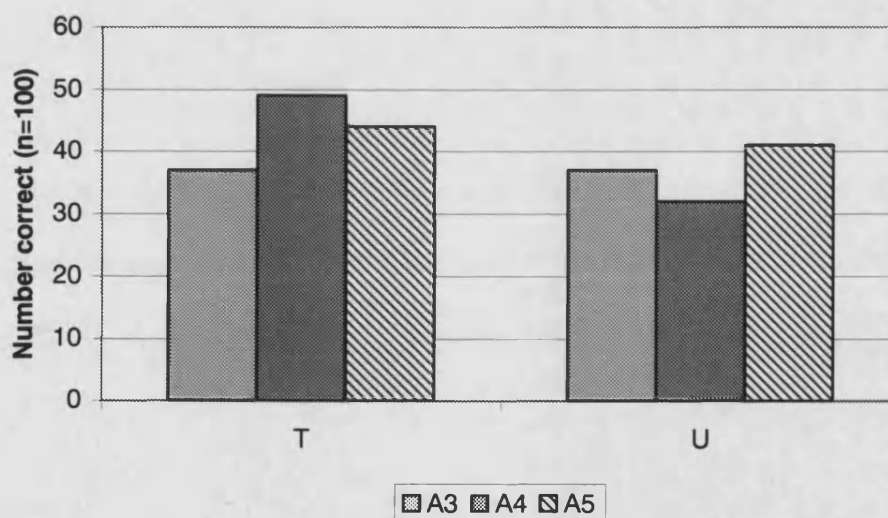


FIGURE 10.8: SC therapy Phase Two treated and untreated sets

10.2.3.2.2 Comparison of therapy effect across four subsets

Analysis of change in the four subsets was carried out to determine where the therapy effect found in Phase Two was located. SC's proportion correct in the four subsets at

assessments one to five are given in Table 10.7, along with significance levels for McNemar tests comparing scores at assessments three and four.

TABLE 10.7: SC scores at assessments one to five for the four subsets of items

	A1	A2	A3	A4	A5	Value of p A3 to A4
tT(n=50)	0.32	0.36	0.36	0.40	0.34	0.38
tU(n=50)	0.34	0.38	0.38	0.38	0.34	0.62
uT(n=50)	0.34	0.38	0.38	0.58	0.54	0.007**
uU(n=50)	0.34	0.36	0.36	0.26	0.48	0.13

A1 to A5 = assessments one to five; tT = treated treated set, tU = treated untreated set, uT = untreated treated set, uU = untreated untreated set. Significance levels are: * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$. Significance levels are for McNemar tests comparing scores at assessments three and four.

The set which was treated in both phases of therapy (tT) improved after Phase Two but this was not significant. The set which was treated in Phase One and left untreated in Phase Two (tU) also did not show any significant change. The set which was treated in Phase Two only (uT) was the only set to improve significantly after Phase Two. The set which was untreated throughout deteriorated after Phase Two but this was not significant. At assessment four there was also a difference between the two untreated sets: those treated in Phase One (tU) maintained their level of performance, whereas those left untreated in both phases deteriorated at assessment four.

There was no significant change in three of the four subsets at assessment five (McNemar tests comparing scores in each subset at assessments four and five: treated Treated: $p = 0.29$, n.s.; treated Untreated: $p = 0.38$, n.s.; untreated Treated: $p = 0.40$, n.s.) However, the untreated Untreated set improved significantly after the period of no intervention (McNemar test: $p = 0.004$). This unexpected improvement means that the other change found in SC's naming (that in the uT set at assessment four) must be viewed with some caution. SC's performance is remarkably consistent prior to assessment four. The change found in the uT set at assessment four may well be due to the therapy, but as substantial change occurred in the uU set at assessment five this cannot be argued unequivocally.

10.2.3.2.3 Analysis of variables across four subsets

As the only effect of therapy occurred in the set which was treated in Phase Two only, and no significant effect of therapy was seen in the other sets, and an unexpected improvement occurred in the uU set at assessment five, key variables were compared to

identify whether any differences existed between sets. SC's naming was predicted by lexical variables and to a lesser extent by animacy. These variables were therefore investigated in order to determine whether differences in their values might have influenced response to therapy. See Table 10.8 for descriptive data.

There was no evidence of any difference between sets in terms of any one of the three variables under consideration. One way ANOVA's conducted on the four sets of scores for each of the three variables were not significant (frequency: $F = 0.68$, $df = 3$, 196 , $p = 0.56$, n.s.; age of acquisition: $F = 1.24$, $df = 3$, 196 , $p = 0.30$, n.s.; familiarity: $F = 0.95$, $df = 3$, 196 , $p = 0.42$, n.s.).

TABLE 10.8: Values of three variables for the four treatment subsets for SC

Subset (n=50)	Frequency	Age of acquisition	Familiarity
tT	1.34 (0.64)	2.71 (0.68)	535 (69)
tU	1.31 (0.49)	2.78 (0.70)	531 (65)
uT	1.28 (0.57)	2.67 (0.53)	516 (61)
uU	1.18 (0.62)	2.91 (0.73)	531 (46)

Figures given are means for the subset for each variable. Standard deviations are in parentheses.

A similar analysis was carried out to investigate whether there were differences across the sets in terms of the animacy of the items. Again no significant difference was found, with the number of inanimate items in each set ranging from 13 to 10 (chi square = 1.63 , $df = 3$, $p = 0.65$, n.s.).

10.2.3.2.4 Summary of SC's response to therapy Phase Two

There was no overall change in the 200 items as a result of Phase Two therapy indicating that there was no overall improvement in word finding. There was a significant change in the set of items treated in Phase Two, but this was offset by a non-significant deterioration in items left untreated in Phase Two.

Of the items treated in Phase Two only those which had not previously received any treatment improved as a result of Phase Two therapy. There was no effect on the items which had already been treated in Phase One - and which had not improved as a result of that therapy - and no effect on untreated items. The isolated effect on the set treated

in Phase Two only cannot be attributed to an underlying variable as there was no difference in terms of key variables across the sets. The set which responded to treatment in Phase Two (uT set) was of lower familiarity than the other treated set (tT) although this was not significant.

Three sets remained stable at follow-up. The set of items which had not received any therapy at all improved after this period of no intervention. Analysis of key variables failed to find a significant difference between this set and the other sets.

As SC failed to respond to Phase One therapy it is unclear why one set of items should respond to Phase Two therapy. He did not respond to cues nor to the related therapy, so the hypothesised mechanism responsible for therapy effects – processing improvements in the mappings between semantics and phonology – could not be responsible. It is possible that repeated use of the target set in more natural communicative situations rendered them easier to access. In addition the set may be ‘easier’ for SC, although this cannot be attributed to any of the variables predicting his naming performance.

10.2.3.3 Participant KR

When the change in scores between assessments three and four was analysed for the total set of 200 items no difference was found (McNemar test: $p = 0.39$). This indicates that Phase Two therapy did not have an overall effect on naming.

10.2.3.3.1 Comparison of treated and untreated sets

KR showed an item-specific response to therapy in Phase One. A comparison of therapy effects for treated versus untreated items for Phase Two showed that therapy led to an improvement for treated items which was not statistically significant (McNemar test: $p = 0.04$), and no significant difference between assessments three and four for those items not treated (McNemar test: $p = 0.13$, n.s.). KR’s performance in the two sets is shown in Figure 10.9. The graph indicates that the treated set improved after Phase Two therapy and continued to improve during the follow-up period. The untreated set showed some deterioration at assessment four, which suggests that the improvement seen in the treated set was at the expense of items in the untreated set.

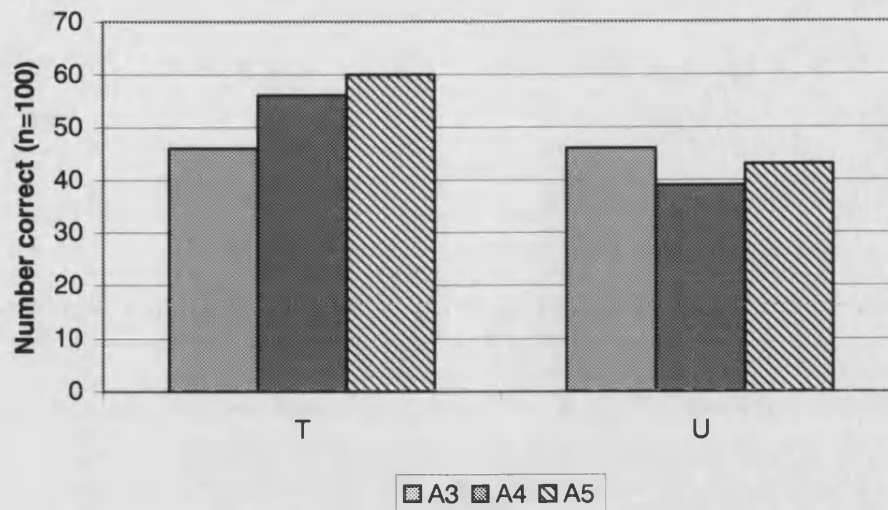


FIGURE 10.9 KR therapy Phase Two treated and untreated sets

10.2.3.3.2 Comparison of therapy effect across four subsets

Analysis of change in the four subsets was carried out to determine where the therapy effect found in Phase Two was located. KR's proportion correct in the four subsets at assessments one to five are given in Table 10.9, along with significance levels for McNemar tests comparing scores at assessments three and four.

TABLE 10.9: KR scores at assessments one to five for the four subsets of items

	A1	A2	A3	A4	A5	Value of p A3 to A4
tT(n=50)	0.40	0.36	0.50	0.46	0.58	0.40
tU(n=50)	0.40	0.36	0.50	0.38	0.50	0.11
uT(n=50)	0.40	0.38	0.42	0.66	0.62	0.001***
uU(n=50)	0.40	0.38	0.42	0.40	0.36	0.50

A1 to A5 = assessments one to five; tT = treated treated set, tU = treated untreated set, uT = untreated treated set, uU = untreated untreated set. Significance levels are: * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$. Significance levels are for McNemar tests comparing scores at assessments three and four.

The set which was treated in both phases of therapy (tT) showed no response to Phase Two therapy. The set which was treated in Phase One and left untreated in Phase Two (tU) showed a deterioration after Phase Two therapy. The set which was treated in Phase Two only (uT) showed a significant improvement after Phase Two. The set which was untreated throughout showed no change.

Comparisons between performance at assessment four and assessment five showed the following results. For all four sets changes were not significant. Results of McNemar tests were: tT: $p = 0.1189$; tU: $p = 0.1051$; uT: $p = 0.3872$; uU: $p = 0.4073$.

In the tT set there was an effect of therapy after Phase One, then instability in the set thereafter. The tU set showed a response to Phase One therapy and again instability thereafter. The uT set showed a stable baseline prior to Phase Two therapy and an effect of Phase Two therapy which is maintained at follow-up. The uU set showed stable performance throughout.

10.2.3.3.3 Summary of KR's response to Phase Two therapy

KR showed no effect overall of the therapy administered in Phase Two when scores across the 200 items were compared at assessments three and four. When the treated and untreated items in Phase Two were compared an item specific effect of therapy was found, with only items treated in Phase Two benefiting, and some deterioration in items not treated in Phase Two. When the four subsets were considered the therapy effect found in the 100 treated items was found to be located specifically in items treated only in Phase Two. Items which had been previously treated in Phase One did not respond to Phase Two therapy. This is similar to SC and PH both of whom showed a tendency for items previously untreated to respond more to therapy in Phase Two.

There was some instability in the two sets treated in Phase One after Phase Two therapy but changes were non-significant. Both deteriorated after Phase Two therapy and both improved at follow-up. Items which were untreated throughout remained at baseline level.

KR had shown some response to therapy in Phase One, which was more marked for the phonological set than for the orthographic set. She is similar to PH in terms of her language profile and one might therefore expect her to follow PH in responding to therapy in Phase Two. This was the case for KR.

10.3 EFFECTS OF THERAPY ON CONVERSATION

In Chapter Six the baseline conversation data for ten people with aphasia was analysed and aspects of conversation were identified which showed satisfactory inter-rater and intra-rater reliability, and test-retest stability. As any change seen in conversation after therapy could be part of the inherent variation present in conversation data, and not therefore attributable to the therapy, it was considered desirable to identify aspects of conversation which are reliably coded by two independent raters, and by the same rater on two occasions, and to identify which aspects are stable over time. This would enable any true effects of therapy on conversation to be identified. The aspects of conversation which showed satisfactory reliability are shown in Table 10.10.

10.3.1 Measuring change after therapy

The reliability and stability data for these aspects of conversation data notwithstanding, it is still possible that for a given individual there is substantial variation over the baseline phase. This would make any change after therapy hard to interpret and certainly one could not conclude a real effect of therapy in such circumstances.

In order to measure any change after therapy the differences between assessment points before and after therapy for each individual were therefore compared to the overall amount of variation seen in the baseline phase for the group of ten people whose data was used in the reliability study (see Chapter Six). The amount of variation in the baseline phase overall was estimated by taking the standard deviation of the differences between scores at assessment one and assessment two. The absolute differences between scores at assessment one and assessment two were taken for each individual, and the mean, standard deviation and range of differences computed⁵². These data give a measure of the amount of variation within the baseline phase for the group of people with aphasia, and are reported in Table 10.11.

The data for the three individuals reported in this thesis were then considered in relation to the data from the baseline for the whole group. Each individual's sets of scores for assessments one to five for draft three of the conversation measure were taken, and the

⁵² For the set of difference scores (between assessment one and assessment two) outliers of more than two standard deviations from the mean were replaced by the group mean (see Tabachnick and Fidell, 2001).

TABLE 10.10 Conversation Measure Draft Three

B	SPEECH UNITS
B1	Person with aphasia
B2	Conversational partner
C	TURN TAKING
C1	Number of turns: Person with aphasia
C2	Number of turns: Conversational partner
C3	Number of substantive turns: Person with aphasia
C4	Number of substantive turns: Conversational partner
C5	Number of minimal turns: Person with aphasia
C6	Number of minimal turns: Conversational partner
D	WORD RETRIEVAL AND SPEECH ERRORS (PERSON WITH APHASIA)
D1	Total number of content words (excluding paraphasias)
D2	Total number of nouns (subset of content words)
D11	Number of filled pauses within person with aphasia's turn
E	REPAIR
E1	Number of instances of repair
F	PROPORTIONAL DATA
F1	Substantive turns / turns
F2	Minimal turns / turns
F3	Content words / speech units
F4	Nouns / speech units
F5	Filled pauses / speech units
F6	Repairs / turns

difference between consecutive assessment points derived. Two forms of comparison were used. Firstly the amount of difference found between two scores for any given individual was compared to the range of difference found in the group baseline data (Table 10.11). This allowed a comparison to be made between a change in scores for an individual, and the range of differences found in the group of people with aphasia. Scores falling outside the range were of interest here. Secondly the difference found between two scores at two consecutive assessment points for an individual on any given variable was divided by the standard deviation of the set of differences found across baseline scores for the group, to give a z score. For example, PH produced 51 nouns at assessment three and 50 nouns at assessment two, giving a difference of one from assessment three to assessment two. This was then divided by the standard deviation of the differences between scores for this variable across the baseline scores. This gave a measure of the amount of change across assessment points, in relation to the change

found in the baseline group data. The latter is thus being treated here as the norm. The analysis conducted here was suggested by M. Szczerbinski (personal communication).

Tables 10.12 to 10.14 give the data from these comparisons for the three people with aphasia. In these tables the raw data from the five assessment points are given, along with the actual differences between any two consecutive assessment points. The final columns show the difference between an individual's two consecutive scores, divided by the standard deviation of the group baseline differences. Any figures in bold type in the columns referring to A2-A1 in Table 10.12 to 10.14 indicate that there is a lack of stability in the baseline and thus any changes in this variable after therapy must be regarded with caution. For the differences, the bold type indicates that the value falls outside the range of the group baseline data (see Table 10.11). For the z scores the bold type indicates a value greater than 1.96: 95% of values fall within 1.96 standard deviations of the mean, with 5% of values falling outside these extremes. Any changes in the columns marked A3-A2 relate to changes after Phase One therapy. For PH and SC changes in the columns marked A4-A3 relate to changes after Phase Two therapy, and changes in the columns marked A5-A4 relate to changes after the period of no intervention. For KR changes in the columns marked A5-A3 relate to changes in the period after Phase Two therapy and the no intervention period combined.

The data from these three tables will be discussed individually with reference to changes shown in bold type. For some of the variables the desirable outcome is an increase in scores, e.g. noun production. For others the desirable outcome is a decrease, e.g. in filled pauses. For other variables this is equivocal e.g. number of turns. An ideal pattern in the data would be no change in the baseline, increases and decreases (as appropriate) after both periods of therapy, and no change after no intervention.

TABLE 10.11: Baseline conversation data for group (n= 10) assessments one and two

		Assessment One		Assessment Two		Difference: assessment one and two			
		Mean A1	St. Dev. A1	Mean A2	St. Dev. A2	Mean	St. Dev.	Min	Max
	Conversation Variable								
B	SPEECH UNITS								
B1	Person with aphasia	321	132	291	162	56.9	34.78	17	129
B2	Conversational partner	319	108	330	113	52.9	44.64	7	122
C	TURN TAKING								
C1	Number of turns: PA	54.5	18.8	45.8	9.9	8.87	5.99	1	15
C2	Number of turns: CP	34.1	12.1	24.9	9.4	8.87	6.92	0	20
C3	Number of substantive turns: PA	42.5	14	35.1	9.6	10.8	5.71	2	17
C4	Number of substantive turns: CP	9.3	8.2	13.9	6.3	6.74	4.96	0	19
C5	Number of minimal turns: PA	8.7	7.4	7.7	7.5	4.56	2.34	0	7
C6	Number of minimal turns: CP	54.5	18.7	45.8	9.9	2.86	1.77	1	6
D	LEXICAL RETRIEVAL								
D1	Total content words	89.4	49.5	75.8	51.4	25.8	17.53	0	54
D2	Total nouns	28.1	13	26.6	17	8.9	6.31	0	20
D11	Filled pauses	16.6	10.9	15.5	11.7	5.7	3.13	0	8
E	REPAIR								
E1	Instances of collaborative repair	7.3	3.7	6.3	3	2.8	1.32	0	4
F	PROPORTIONAL DATA								
F1	Substantive turns / turns	0.65	0.23	0.52	0.16	0.12	0.07	0.01	0.23
F2	Minimal turns / turns	0.14	0.08	0.30	0.13	0.14	0.07	0.00	0.23
F3	Content words / speech units	0.26	0.08	0.24	0.08	0.03	0.02	0.01	0.06
F4	Nouns / speech units	0.09	0.04	0.09	0.03	0.02	0.01	0.00	0.04
F5	Filled pauses / speech units	0.07	0.06	0.06	0.05	0.02	0.01	0.00	0.03
F6	Instances of repair / total turns	0.07	0.03	0.07	0.03	0.02	0.01	0.00	0.04

TABLE 10.12 Pre and post therapy conversation data for PH

		Raw scores					Differences				Z scores in relation to group baseline score			
		A1	A2	A3	A4	A5	A2-A1	A3-A2	A4-A3	A5-A4	A2-A1	A3-A2	A4-A3	A5-A4
B	SPEECH UNITS													
B1	Number of speech units (PA)	566	544	372	486	464	-22	-172	114	-22	-0.63	-4.95	3.28	-0.63
B2	Number of speech units (CP)	302	284	349	336	371	-18	65	-13	35	-0.40	1.46	-0.29	0.78
C	TURN TAKING													
C1	Total number of turns (PA)	72	71	82	65	70	-1	11	-17	5	-0.17	1.84	-2.84	0.83
C2	Total number of turns (CP)	69	66	86	65	73	-3	20	-21	8	-0.43	2.89	-3.03	1.16
C3	Number of substantive turns (PA)	49	42	60	45	47	-7	18	-15	2	-1.23	3.15	-2.63	0.35
C4	Number of substantive turns (CP)	46	39	47	44	47	-7	8	-3	3	-1.41	1.61	-0.60	0.60
C5	Number of minimal turns (PA)	15	13	9	14	13	-2	-4	5	-1	-0.85	-1.71	2.14	-0.43
C6	Number of minimal turns (CP)	20	21	22	19	13	1	1	-3	-6	0.56	0.56	-1.69	-3.39
D	LEXICAL RETRIEVAL													
D1	Number of content words (PA)	165	141	125	123	122	-24	-16	-2	-1	-1.37	-0.91	-0.11	-0.06
D2	Number of nouns (PA)	30	50	51	18	29	20	1	-33	11	3.17	0.16	-5.23	1.74
D11	Number of filled pauses (PA)	7	15	11	0	6	8	-4	-11	6	2.56	-1.28	-3.51	1.92
E	REPAIR													
E1	Number of repairs	7	7	8	6	4	0	1	-2	-2	0.00	0.76	-1.52	-1.52
F	PROPORTIONAL DATA													
F1	Substantive turns / turns	0.68	0.59	0.73	0.69	0.67	-0.09	0.14	-0.04	-0.02	-1.31	2.04	-0.58	-0.29
F2	Minimal turns / turns	0.21	0.18	0.11	0.22	0.19	-0.03	-0.07	0.11	-0.03	-0.41	-0.96	1.51	-0.41
F3	Content words / speech units	0.29	0.26	0.34	0.25	0.26	-0.03	0.08	-0.09	0.01	-1.80	4.81	-5.41	0.60
F4	Nouns / speech units	0.05	0.09	0.14	0.04	0.06	0.04	0.05	-0.1	0.02	3.41	4.27	-8.53	1.71
F5	Filled pauses / speech units	0.01	0.03	0.03	0	0.01	0.02	0	-0.03	0.01	1.73	0.00	-2.59	0.86
F6	Repairs / turns	0.05	0.05	0.05	0.05	0.03	0	0	0	-0.02	0.00	0.00	0.00	-1.81

Raw scores, differences between two sets of raw scores, and differences between two sets of raw scores divided by the standard deviation of difference scores for the group of 10 people with aphasia at baseline. Figures in bold in the Differences columns are those which fall outside the range of difference scores of the baseline group. Figures in bold in the z scores columns are those whose value is greater than 1.96. Negative values in the differences and z scores columns denote a decrease in score from one assessment point to the immediately following assessment point. Positive values denote increases in scores.

TABLE 10.13: Pre and post therapy conversation data for SC

		Raw scores					Differences				Z scores in relation to group baseline score			
		A1	A2	A3	A4	A5	A2-A1	A3-A2	A4-A3	A5-A4	A2-A1	A3-A2	A4-A3	A5-A4
B	SPEECH UNITS													
B1	Number of speech units (PA)	401	478	397	487	510	77	-81	90	23	2.21	-2.33	2.59	0.66
B2	Number of speech units (CP)	323	201	257	272	335	-122	56	15	63	-2.73	1.25	0.34	1.41
C	TURN TAKING													
C1	Total number of turns (PA)	66	51	45	47	39	-15	-6	2	-8	-2.50	-1.00	0.33	-1.34
C2	Total number of turns (CP)	66	46	47	46	39	-20	1	-1	-7	-2.89	0.14	-0.14	-1.01
C3	Number of substantive turns (PA)	49	32	40	39	30	-17	8	-1	-9	-2.98	1.40	-0.18	-1.58
C4	Number of substantive turns (CP)	48	23	24	29	28	-25	1	5	-1	-5.04	0.20	1.01	-0.20
C5	Number of minimal turns (PA)	5	12	3	5	7	7	-9	2	2	2.99	-3.85	0.85	0.85
C6	Number of minimal turns (CP)	15	19	22	18	10	4	3	-4	-8	2.26	1.69	-2.26	-4.52
D	LEXICAL RETRIEVAL													
D1	Number of content words (PA)	101	125	141	131	126	24	16	-10	-5	1.37	0.91	-0.57	-0.29
D2	Number of nouns (PA)	33	46	31	41	33	13	-15	10	-8	2.06	-2.38	1.58	-1.27
D11	Number of filled pauses (PA)	6	3	13	9	9	-3	10	-4	0	-0.96	3.19	-1.28	0.00
E	REPAIR													
E1	Number of repairs	1	4	3	3	2	3	-1	0	-1	2.27	-0.76	0.00	-0.76
F	PROPORTIONAL DATA													
F1	Substantive turns / turns	0.74	0.63	0.89	0.83	0.77	-0.11	0.26	-0.06	-0.06	-1.60	3.78	-0.87	-0.87
F2	Minimal turns / turns	0.08	0.24	0.07	0.11	0.18	0.16	-0.17	0.04	0.07	2.19	-2.33	0.55	0.96
F3	Content words / speech units	0.25	0.26	0.36	0.27	0.25	0.01	0.1	-0.09	-0.02	0.60	6.01	-5.41	-1.20
F4	Nouns / speech units	0.08	0.1	0.08	0.08	0.06	0.02	-0.02	0	-0.02	1.71	-1.71	0.00	-1.71
F5	Filled pauses / speech units	0.01	0.01	0.03	0.02	0.02	-0.01	0.03	-0.01	0.00	-0.75	2.28	-1.23	-0.07
F6	Repairs / turns	0.01	0.04	0.03	0.03	0.03	0.03	-0.01	0	0	2.71	-0.90	0.00	0.00

Raw scores, differences between two sets of raw scores, and differences between two sets of raw scores divided by the standard deviation of difference scores for the group of 10 people with aphasia at baseline. Figures in bold in the Differences columns are those which fall outside the range of difference scores of the baseline group. Figures in bold in the z scores columns are those whose value is greater than 1.96. Negative values in the differences and z scores columns denote a decrease in score from one assessment point to the immediately following assessment point. Positive values denote increases in scores.

TABLE 10.14 Pre and post therapy conversation data for KR

		Raw scores					Differences				Z scores in relation to group baseline score			
		A1	A2	A3	A4	A5	A2-A1	A3-A2	A4-A3	A5-A3	A2-A1	A3-A2	A4-A3	A5-A3
B	SPEECH UNITS													
B1	Number of speech units (PA)	246	299	309	-	199	53	10	-	-110	1.52	0.29	-	-3.16
B2	Number of speech units (CP)	340	350	461	-	695	10	111	-	234	0.22	2.49	-	5.24
C	TURN TAKING													
C1	Total number of turns (PA)	69	54	94	-	98	-15	40	-	4	-2.50	6.68	-	0.67
C2	Total number of turns (CP)	68	55	90	-	96	-13	35	-	6	-1.88	5.06	-	0.87
C3	Number of substantive turns (PA)	46	32	49	-	30	-14	17	-	-19	-2.45	2.98	-	-3.33
C4	Number of substantive turns (CP)	46	42	53	-	71	-4	11	-	18	-0.81	2.22	-	3.63
C5	Number of minimal turns (PA)	13	20	24	-	58	7	4	-	34	2.99	1.71	-	14.53
C6	Number of minimal turns (CP)	22	11	25	-	23	-11	14	-	-2	-6.21	7.91	-	-1.13
D	LEXICAL RETRIEVAL													
D1	Number of content words (PA)	76	98	83	-	50	22	-15	-	-33	1.25	-0.86	-	-1.88
D2	Number of nouns (PA)	46	32	34	-	18	-14	2	-	-16	-2.22	0.32	-	-2.54
D11	Number of filled pauses (PA)	37	37	29	-	9	0	-8	-	-20	0.00	-2.56	-	-6.39
E	REPAIR													
E1	Number of repairs	8	4	3	-	2	-4	-1	-	-1	-3.03	-0.76	-	-0.76
F	PROPORTIONAL DATA													
F1	Substantive turns / turns	0.67	0.59	0.52	-	0.31	-0.08	-0.07	-	-0.22	-1.16	-1.00	-	-3.13
F2	Minimal turns / turns	0.19	0.37	0.26	-	0.59	0.18	-0.11	-	0.34	2.47	-1.57	-	4.61
F3	Content words / speech units	0.31	0.33	0.27	-	0.25	0.02	-0.06	-	-0.02	1.20	-3.69	-	-1.04
F4	Nouns / speech units	0.19	0.11	0.11	-	0.09	-0.08	0.00	-	-0.02	-6.83	0.00	-	-1.67
F5	Filled pauses / speech units	0.15	0.12	0.09	-	0.05	-0.03	-0.03	-	-0.05	-2.30	-2.58	-	-4.20
F6	Repairs / turns	0.06	0.04	0.02	-	0.01	-0.02	-0.02	-	-0.01	-1.81	-2.14	-	-0.54

Raw scores, differences between two sets of raw scores, and differences between two sets of raw scores divided by the standard deviation of difference scores for the group of 10 people with aphasia at baseline. Figures in bold in the Differences columns are those which fall outside the range of difference scores of the baseline group. Figures in bold in the z scores columns are those whose value is greater than 1.96. Negative values in the differences and z scores columns denote a decrease in score from one assessment point to the immediately following assessment point. Positive values denote increases in scores. For KR there are no data for assessment four. The final column of differences and z scores relate to the difference between assessments five and three.

10.3.2 Conversation outcome for PH

10.3.2.1 Speech units

The number of speech units produced by PH is stable at the baseline, decreases after Phase One, and increases after Phase Two, and is stable at follow-up. There are no notable changes in the number of speech units for the conversation partner.

10.3.2.2 Turn taking

The total number of turns produced by PH is stable at baseline, shows no notable change after Phase One, but decreases after Phase Two. The total number of turns produced by PH's conversation partner is stable at baseline, increases after Phase One therapy, decreases after Phase Two, and is stable at follow-up.

The number of substantive turns PH produces is stable at baseline, increases after Phase One therapy, decreases after Phase Two therapy, and is stable at follow-up. The number of substantive turns PH's partner produces shows no notable change throughout. The number of minimal turns PH produces is stable at baseline, shows no change after Phase One therapy, increases after Phase Two therapy (which is undesirable), and is stable at follow-up. The conversation partner's minimal turns are stable at baseline and after both phases of therapy, but decrease after no intervention.

10.3.2.3 Lexical retrieval and repair

PH's production of content words shows no notable changes. Her production of nouns increases in the baseline from assessment one to two, so any data relating to this must be viewed cautiously. Production of nouns decreases after Phase Two therapy. Filled pauses are also unstable at baseline, so subsequent changes are hard to interpret. After Phase One and after Phase Two there is a decrease in filled pauses. There are no notable changes in the number of repairs.

10.3.2.4 Proportional data

The proportion of substantive turns shows a stable baseline and an increase after Phase One therapy, no change after Phase Two therapy and no change after no intervention.

The proportion of minimal turns shows no notable changes. The proportion of content words shows a stable baseline, an increase after Phase One therapy, a decrease after Phase Two therapy, and no change after no intervention. The proportion of nouns shows an unstable baseline, an increase after Phase One therapy, a decrease after Phase Two therapy, and no change after no intervention. The proportion of filled pauses shows a stable baseline, no change after Phase One therapy, a decrease after Phase Two therapy, and no change after no intervention. The proportion of repair shows no notable change throughout.

10.3.2.5 Summary of PH's conversation data

Most of PH's baseline data are stable meaning that it is more likely that changes after therapy are real effects of therapy. There is a pattern in the data which suggests that therapy Phase One had a positive impact on conversation. After Phase One PH produces more substantive turns, and a higher proportion of substantive turns, of content words and of nouns, although the unstable baseline in the latter case make these data unreliable. After Phase Two therapy she produced fewer filled pauses (also with an unstable baseline), and a smaller proportion of filled pauses, but also fewer turns overall, fewer substantive turns, more minimal turns, fewer nouns, and a smaller proportion of content words and of nouns (but see comments regarding baseline). There were few changes after the period of no intervention, apart from the decrease seen in the number of minimal turns PH's partner produced.

10.3.3 Conversation outcome for SC

10.3.3.1 Speech units

SC's production of speech units showed an unstable baseline, a decrease after Phase One therapy, an increase after Phase Two therapy, and no change at follow-up. His partner's production of speech units showed an unstable baseline, and no further changes.

10.3.3.2 Turn taking

Both SC's and his partner's production of turns and their production of substantive turns showed an unstable baseline, and no further changes. SC's production of minimal turns

showed an unstable baseline, a decrease after therapy Phase One, and no further changes. SC's partner's production of minimal turns showed an unstable baseline, no change after therapy Phase One, a decrease after therapy Phase Two and a further decrease at follow-up.

10.3.3.3 Lexical retrieval and repair

SC's production of content words showed no notable change throughout. His production of nouns showed an unstable baseline, a decrease after therapy Phase One, and no further changes. Filled pauses showed a stable baseline, an increase after therapy Phase One, and no further changes. The instances of repair showed an unstable baseline, and no further changes.

10.3.3.4 Proportional data

The proportion of substantive turns showed a stable baseline, an increase after therapy Phase One, and no further changes. The proportion of minimal turns showed an unstable baseline, and a decrease after therapy Phase One, and no further changes. The proportion of content words showed a stable baseline, an increase after therapy Phase One, a decrease after therapy Phase Two, and no further change. This is similar to the pattern found for PH. Unlike for PH however there is no other evidence to support a claim of improvement after therapy. The proportion of nouns showed no notable changes throughout. The proportion of filled pauses showed an unstable baseline, an increase after therapy Phase One and no further changes. The proportion of repair showed an unstable baseline and no further changes.

10.3.3.5 Summary of SC's conversation data

Most of the variables considered here showed an unstable baseline, thus any changes witnessed thereafter cannot be confidently attributed to the therapy. After therapy Phase One SC produced fewer minimal turns, a smaller proportion of minimal turns, and a larger proportion of substantive turns and content words. He also produced fewer speech units, fewer nouns, more filled pauses, and a larger proportion of filled pauses.

After therapy Phase Two SC produced more speech units, and his partner produced fewer minimal turns. His proportion of content words decreased after Phase Two. At

follow-up there were no changes to SC's data, although his partner's production of minimal turns decreased.

Given the unstable baseline scores, it is difficult to make strong conclusions about SC's conversation data. It is important to note however that he made no change in naming overall as a result of the two therapies, therefore it is unlikely that there would be a significant difference in word retrieval in conversation after therapy.

10.3.4 Conversation outcome for KR

It is important to note that KR's final conversation was unlike the previous three in that she took a relatively passive part. KR's husband introduced a topic, and KR's main contribution was to provide minimal turns to support his exposition. It may be that this topic was too difficult in terms of vocabulary for her to do more than provide turns which handed the floor back to her husband. This variation in conversations is one reason why this form of data is so difficult to measure and to interpret.

10.3.4.1 Speech units

KR's production of speech units showed a stable baseline, no change after Phase One therapy, and a decrease after combined period of Phase Two and follow-up. Her partner's speech unit production showed a stable baseline, an increase after Phase One, and an increase was found at assessment five.

10.3.4.2 Turn taking

KR's production of turns showed an unstable baseline, an increase after therapy Phase One, and no further change. Her partner's turn production showed a stable baseline, an increase after therapy Phase One, and no further change. KR's substantive turn production showed an unstable baseline, an increase after therapy Phase One, and a decrease at assessment five. Her partner's substantive turn production showed a stable baseline, an increase after therapy Phase One, and an increase at assessment five. KR's minimal turn production showed an unstable baseline, no change after Phase One, and an increase at assessment five. Her partner's minimal turn production showed an unstable baseline, an increase after Phase One and no further change.

10.3.4.3 Lexical retrieval and repair

Content word production showed a stable baseline, and no changes. Noun production showed an unstable baseline, no change after Phase One, and a change at assessment five. Filled pauses showed a stable baseline, a decrease after Phase One, and a decrease at follow-up assessment five. Repair showed an unstable baseline and no further changes.

10.3.4.5 Proportional data

The proportion of substantive turns produced by KR showed a stable baseline, no change after Phase One and a decrease at assessment five. The proportion of minimal turns showed an unstable baseline, no change after Phase One, and an increase at assessment five. The proportion of content words showed a stable baseline, a decrease after Phase One, and a further decrease at assessment five. The proportion of nouns showed an unstable baseline, no change after Phase One, and a decrease at assessment five. The proportion of filled pauses showed an unstable baseline, a decrease after Phase One, and a further decrease at assessment five. The proportion of repairs showed a stable baseline, a decrease after Phase One and no further change.

10.3.4.6 Summary of KR's conversation data

A large number of variables showed unstable baselines and therefore any data relating to possible therapy effects is to be viewed with caution. After Phase One therapy KR and her partner produced more turns and more substantive turns. KR produced fewer filled pauses, a smaller proportion of filled pauses, and there was a smaller proportion of repair. In addition KR's partner produced more speech units and more minimal turns. Also after Phase One KR produced a smaller proportion of content words.

After the combined period of Phase Two and a period of no intervention, KR produced fewer filled pauses, and a smaller proportion of filled pauses. She also produced fewer speech units, fewer substantive turns, more minimal turns, fewer nouns, a smaller proportion of substantive turns, and a larger proportion of minimal turns. Her partner in contrast produced more speech units, and more substantive turns.

10.3.5 Summary of conversation data

In Chapter Six conversation variables were identified which were thought to show good test-retest stability. This was measured in terms of the correlation coefficient. Using the standard deviation of the group baseline differences to compute z scores, the conversation data from the three people with aphasia was examined in terms of the individual stability of the baseline, and the amount of change shown after both phases of therapy and after the follow-up period.

Importantly a number of variables showed unstable individual baselines. This makes any further conclusions about the data hard to draw. There are some indications however that therapy had an effect on conversation for some aspects of conversation for some of the individuals involved. These will be discussed in the next chapter. There are also strong indications that there is significant variation in conversation data across assessment times, making its use as an outcome measure problematic.

10.4 LANGUAGE CONTROL TASKS

A set of four language control tasks were repeated at each of the five assessment points throughout the study. These were: written sentence comprehension, reading aloud a set of 52 words, reading aloud a set of 26 non-words, and a measure of short term memory which involved listening to a list of picture names then pointing to each in the designated order. Where a participant was at floor on a task this was administered in the auditory rather than the written modality. Thus for SC, whose reading was severely compromised, sentence comprehension was administered via the auditory route, and repetition rather than reading aloud of words and non-words was carried out. KR was at floor on reading non-words so she also completed repetition of this set.

10.4.1 Results

10.4.1.1 Results for PH

PH only completed four of the five assessments, as she was unavailable for assessment two (see Table 10.15). It is not possible therefore to conclude anything about her baseline. Her overall performance in this assessment shows a significant trend towards improvement when all the data is considered (Jonckheere Trend Test: $z = 2.80$, one

tailed $p = 0.0026$). When assessment one is discarded this is not significant (Jonckheere Trend Test: $z = 0.14$, one tailed $p = 0.444$, n.s.).

TABLE 10.15 PH language control tasks

PH	A1	A2	A3	A4	A5
Written sentence comprehension (n=16)	3	-	11	10	12
Reading aloud words (n=52)	50	-	52	48	49
Reading aloud non words (n=26)	9	-	14	17	12
Short term memory span	3.5	-	4.5	4.1	3.9

Visual analysis of PH's scores on reading words aloud showed no significant change, but she is near ceiling on this task and therefore the results are not informative. Reading aloud of non-words improved from assessment one to assessment three and from three to four, then deteriorated at assessment five. There was a significant trend towards improvement up to assessment four (Jonckheere Trend Test: $z = 2.07$, one tailed $p = 0.0193$). When the data from assessment five is included this is no longer significant however (Jonckheere Trend Test: $z = 0.96$, one tailed $p = 0.169$). Short term memory span showed an increase at assessment three then little change thereafter.

As there is inadequate baseline data it is difficult to draw firm conclusions from the data. Nevertheless there appears to be an overall improvement in two tasks (written sentence comprehension and reading aloud non-words). For written sentence comprehension the improvement appears to occur between assessment one and the subsequent assessments, suggesting that therapy Phase One was the significant factor here. For non-word reading there is an improvement at assessment three and a smaller improvement at assessment four, but this is not sustained at assessment five. In both cases there are two possible explanations to be considered. The first is that the data represent a real improvement in performance, and that this is linked to therapy Phase One. This is a counter-intuitive proposition as therapy targeted a separate and discrete element of processing. The second explanation proposes that PH grew accustomed to the test situation and to the researcher and thus later test results better reflect her actual level of performance. Without the second baseline point of assessment this point is impossible to argue, but remains a feasible hypothesis.

10.4.1.2 Results for SC

SC was unable to carry out tasks administered via the written modality and so auditory input was used. The results are shown in Table 10.16. SC showed minimal variation across assessment times for all four tests. There is some evidence of improvement in repetition of words at assessment three. Statistical analysis of change in scores from assessment one to two to three is not significant

TABLE 10.16 SC language control tasks

SC	A1	A2	A3	A4	A5
Auditory sentence comprehension (n=16)	10	12	13	12	12
Repetition words (n=52)	31	28	35	31	34
Repetition non words (n=26)	7	12	7	8	4
Short term memory span	2.7	2.7	2.9	2.9	2.9

however (Jonckheere Trend Test, $z = 0.7$, $p = 0.2423$, n.s.). There is an unstable baseline in SC's repetition of non-words, which improves from 7/26 at assessment one to 12/26 at assessment two. McNemar chi square comparison of the two sets of scores was not significant however ($p = 0.09$, n.s.). Visual analysis of scores on the test of short term memory span indicate that although this improved slightly at assessment three and maintained this level, this improvement was small and therefore not indicative of overall change in this function.

Although there is some variation in the baseline for non word repetition, and some evidence of improvement at assessment three in word reading, overall SC produced a stable performance in all four tests across the five assessment points. There is no evidence from the results of these tests therefore of an effect of therapy on language functions not treated directly in the therapy. SC showed no response to Phase One therapy, and a response restricted to one treated set in Phase Two therapy, thus it would be surprising to find change in other language functions not directly related to the function targeted in the two therapies.

10.4.1.3 Results for KR

KR was unable to complete reading aloud of non-words and she completed repetition of non-words instead. Visual analysis of all four sets of scores suggests that there are

stable baselines in KR's scores, and no indication of significant change after therapy. No statistical tests were carried out on the data as there is no evidence of change.

TABLE 10.17 KR language control tasks

KR	A1	A2	A3	A4	A5
Written sentence comprehension (n=16)	9	9	7	8	11
Reading aloud (n=52)	35	35	32	31	30
Repetition non words (n=26)	18	15	16	17	17
Short term memory span	2.1	2.5	2.5	2.7	2.3

The scores from KR's language control tasks suggest that the effects of the therapies conducted in this study were restricted to the treated language function, i.e. word-finding. None of the language functions assessed as the language control tasks showed any improvement. It is therefore safe to conclude that any effects of therapy found in the other outcome measures were related to the therapy itself and not due to an overall improvement in language processing or a generalised improvement in another related mental capacity such as attention.

10.5 SUMMARY OF CHAPTER TEN

Therapy outcomes in terms of picture naming and conversation have been described for the three participants. Overall the lexical therapy involving phonological and orthographic cues was effective for two of the three participants. The interactional therapy was effective for all three participants, at least for the treated subsets, and in particular for the subset not seen in therapy Phase One. The effects of therapy on conversation are harder to interpret owing to unstable baselines. There are some signs that some aspects may have improved for at least one individual. Issues relating to this will be discussed in full in Chapter Eleven.

CHAPTER ELEVEN: GENERAL DISCUSSION

11.0 INTRODUCTION

This thesis reports the results of investigations into three related areas of aphasia therapy. In one an attempt has been made to devise a quantitative measure of word retrieval in conversation. The rationale for this was that therapy effects have been measured routinely in terms of gains in picture naming, with little evidence accruing that therapy may also benefit everyday interaction. There is a pressing need for this type of data in order to be sure that therapy meets the needs of people with aphasia. In addition evidence of positive functional outcomes is becoming more urgently required in order to secure funding for therapy services. At present there is no existing assessment which addresses this concern directly. The second broad area of investigation involved in-depth assessment of three people with aphasia, using many published, standardised tests, but also comparing single word production in test conditions, with single word production in conversation. In this section participants' response to facilitatory cues was also assessed. The third area compared the effects of two different but related forms of therapy for word finding in the three participants with aphasia. All three participants underwent both therapies. Therapy outcomes were measured by picture naming and in conversation using the quantitative measure.

The main hypotheses outlined at the beginning of the thesis were that lexical therapy (Phase One) would be most effective for participants presenting with a deficit in mapping between good semantic and phonological representations. It was also predicted that this form of therapy would have item-specific effects, with no gains found in items left untreated. A logical step from this latter point is that as relatively few items underwent therapy, improvements found in the production of these few items would not affect overall word finding performance in conversation, thus no change was expected in the quantitative assessment of conversation.

It was hypothesised that the interactive therapy (Phase Two) would be effective for a broader range of people with aphasia, incorporating as it does semantic, syntactic and phonological components. As the therapy targeted a specific set of treated items it was predicted that these would improve as a result of therapy. As the therapy involved free selection of words in interaction, it was also predicted that the effect of therapy would

generalise to other words. As a result of this general effect on word finding it was hypothesised that this form of therapy would also lead to improved word finding in conversation.

The discussion is organised around the three main areas outlined above. In the first section the conversation measure is discussed, and future extensions of this work are proposed. In the second section the three participants' aphasia profiles are reviewed. The ability of models of single word production to account for the three profiles is analysed. Issues around the assessment of aphasia are also reviewed in this section. In the third section the results of the two therapies for the three individuals with aphasia are discussed with reference to their overall language profile and in particular their single word production. The relationship between therapy and facilitation is explored, and the outcomes of therapy in terms of changes in the conversation measure reported.

11.1 THE CONVERSATION MEASURE

11.1.1 Applications of the measure

The measure developed here addresses everyday communication directly. As such it provides a major step forward in terms of the assessment of aphasia. It allows researchers and clinicians to compare a person's performance on assessment tasks such as picture naming with performance on key aspects of conversation. If used with adequate baseline data it allows the measurement of change after therapy. Although developed with people with aphasia in mind, and using data uniquely from this group, the measure also has applications with other communication-impaired populations. One example would be its use with children with developmental language disorders. The comparison of children's language test results with their performance on the specific aspects of conversation quantified in the measure would add to investigations in this area of research. Additional applications include its use with other language-impaired adults such as those with progressive neurological disease. With minor modifications the measure could also be used with adults and children with impaired speech, such as dysarthria, or fluency disorders.

In the discussion which follows two crucial aspects of the measure are scrutinised. These relate to the reliability of the measure and its clinical usefulness. The two aspects

are related and are discussed with regard to the development of the measure. Future directions for this research are outlined.

11.1.2 Selection of variables

The variables included in this measure were selected from previous reports, notably from Crockford and Lesser's (1994) similar work, from Schlenk, Huber and Willmes' (1987) definitions of trouble indicating behaviours, from the literature on conversation analysis, and from studies of the use of this with people with aphasia. There are alternatives to the variables selected here however, and to the definitions accorded to particular variables, notably turn types.

Other variables which could have been part of the measure include: adjacency pairs, such as question answer sequences, and in particular people with aphasia's awareness of the turn-taking demands these place; and self initiated self repair. The latter is the commonest form of repair in non-aphasic speakers' conversation. By not incorporating this as a variable the measure is unable to demonstrate how prevalent this form of repair is in people with aphasia. The rationale for the exclusion of these two variables is as follows: adjacency pairs constitute a further analysis of turn-taking, and the measure already included two (substantive and minimal turns) and, in the original version, in line with Crockford and Lesser's (1994) paper, a third (initiations). For reasons of parsimony and for ease of clinical use these two forms of analysis of turn-taking were judged to be adequate. Moreover, published accounts of turn-taking in aphasia (e.g. Lesser and Algar, 1993) indicate that the production of meaningful turns which contribute new semantic information to the conversation may be dispreferred by the person with aphasia over minimal turns, which serve to hand the floor back to the other speaker, thereby lifting the conversational burden from the person with aphasia. It was this aspect of conversational turn-taking that the measure was designed to capture.

Defining what constitutes substantive and minimal turns proved important in this work. Other reports of similar attempts to quantify conversation (e.g. Comrie et al, 2000) define their turns differently, and in the field of discourse analysis minimal speech units such as *um*, *er*, are treated as insignificant (e.g. Ulatowska et al, 1992). For the purposes of this measure substantive turns were defined as turns containing at least one content word where a content word is a noun, verb (excluding modal verbs), adjective, adverb

ending in *-ly*, or a numeral, in line with Bird and Franklin's (1996) criteria. There are problems with this definition however. For example, on occasions when a person with aphasia produces as their next turn a content word which is a repetition of a word produced by the previous speaker in the immediately preceding turn. In this instance, the turn is coded a substantive turn, yet it does not add new information to the conversation or move the conversation on, and acts more like a minimal speech unit in communicating to the first speaker that the person with aphasia is still active and engaged in the conversation. Subtle aspects of intonation could also distinguish a mere repetition of a content word, which should rightly be classified as a minimal turn, from a repetition with a differing intonation contour, such as rising intonation to signal a question. In this instance this might be better coded as a substantive turn as it offers new information about the speaker's viewpoint and calls for a response.

Similarly with minimal turns, both the grammatical class of the items and the function within the turn-taking routine needs to be analysed. For a large number of minimal speech units (*um*, *er*, *oh dear* etc.) the coding of these as such was not in dispute. Problems arose with this category where a person with aphasia produced a phrase such as 'I don't know' or 'I know' which contains a content word and therefore could be classed as a substantive turn, but whose function in the conversation was minimal in that it did not add new information, and served only to hand the floor back. A further decision was made to exclude minimal speech units which formed part of a repair trajectory from the category of minimal turns. For adequate inter-rater and intra-rater reliability to be found on minimal turns, it was then important that both raters, or one rater on two separate occasions, made the same judgements about which were involved in a repair. This was not always the case.

In devising a measure such as this decisions regarding exclusion have to be made, and it is sensible to include aspects of conversation which are known to be problematic from previous studies of aphasia. This means however that areas of difficulty not hitherto identified will not be picked up by the measure. In other words this assessment is not data driven. In addition although the majority of published reports in this area fail to do this, definitions for particular variables are necessary as subjectivity is high in this form of analysis. This point is worth making as variables such as semantic errors, which one would assume to be low subjectivity items, showed extremely poor inter-rater and intra-rater reliability. In terms of turn types, it is safe to conclude from the work carried out

here that, although there are problems with the definitions devised for this measure, that defining substantive turns allowed good reliability. The problems in reliability found in minimal turns could be addressed by placing a greater emphasis on the interaction and less on the actual turn content. This highlights one of the dangers of this form of analysis as it risks taking turns out of their interactional context thereby depriving them of their overall meaning.

In summary the variables included in this measure address aphasic conversation by concentrating on turn types and lexical retrieval. Difficulties in lexical retrieval are known to affect turn-taking quality in aphasia (Perkins et al, 1999), thus by measuring both these aspects the assessment addresses directly one aspect of language processing which is likely to be extremely problematic for people with aphasia. In addition this aspect can also be measured in picture naming, and it was the focus of the two therapy regimes conducted in this study. The study overall therefore provides a coherent account of word finding and its effects in different speech activities.

11.1.3 Sampling issues

11.1.3.1 Conversation partners

The conversation data sampled consisted of five minute conversations recorded between a person with aphasia and a significant other, usually a spouse or another relative, but in some cases a friend. Participants were requested to nominate a conversation partner, and that partner had to be available for each assessment point throughout the study. This often amounted to several months of involvement.

The requirement to have a consistent conversation partner necessarily excluded some people with aphasia, who had no regular contact with a particular individual. Only people who had regular interactions with other people were therefore included in the study. This means that the group sampled here are from one subset of people with aphasia. Those who are more socially isolated, who are participating less in conversation on a daily basis, and whose language and interaction skills may be less stimulated, were not included in the study. An alternative to the sampling method used here, which may get around this problem, is to gather data from conversations between people with aphasia and the researcher or therapist, or between the person with aphasia

and any other lay person. The first alternative may result in a different quality of conversation with issues of power and authority intruding, and a possible interview style of interaction emerging. The second alternative introduces another source of variation. Given the variation across assessment times in conversations between the same two people shown in this study it is likely that even more variables would be unstable across test times. Neither of these two are ideal solutions but they are possible alternatives to the design used here and may be more viable clinically.

11.1.3.2 Sample length

The sample recorded was ten to 15 minutes in length, and the middle five minutes were analysed. For some people with aphasia this was too long and they were unable to produce the required amount of data. They were thus excluded from the study. Most people were however able to produce this amount on each occasion. For reasons relating to reliability (see 11.1.3 below) this sample may be too short. In terms of clinical use this is probably the maximum length that can be tolerated if a full analysis of the sample is to be conducted. If only partial analysis is to be carried out (e.g. counting collaborative repairs only) then a longer sample is possible. Recording a conversation introduces inhibitions and a necessity to keep talking which is difficult for people with aphasia. They therefore differ as to how much of this form of activity they can tolerate. In the clinical setting individual differences would need to be taken into account.

11.1.3.3 Topic constraint

Participants were asked to produce a natural conversation, and to talk about anything they wanted to. This necessarily introduced a source of variation into the data. The problems this might engender were most evident with KR's conversations. In the first three the topic centred around events around the home and family, to which KR could contribute significantly. In the last conversation her husband talked about his university career, a topic to which KR found it difficult to contribute.

Alternatives might be to select one topic and ask participants to talk about this on each occasion, for example family holidays, family members, and so on. This would have the advantage of restricting the topic to an area in which the person with aphasia felt comfortable, while not overly restricting the conversation. Ramsberger and Rende

(2002) restricted the topic in their investigation of transactional success in aphasia by asking people with aphasia to watch a video of a well-known television comedy, and then recount the story to another person. One problem with the measurement of therapy gains in conversation lies in the fact that relatively few words undergo therapy, and it is very unlikely that in any given conversation a significant number of those particular vocabulary items would be required. An alternative would be to target a particular semantic category, and ask participants to conduct conversations around that category. Another possibility would be to use the initial conversations between a person with aphasia and their partner to identify targets for therapy. This would involve an analysis of the types of topics under discussion, and the words causing problems in retrieval. Therapy targeted at these words would ideally lead to gains in picture naming, and increased word retrieval in conversation.

11.1.3.4 Tape recorded data

The rationale for using tape-recording instead of video was that this measure is designed to capture word finding deficits and their effect on conversation, and successful or unsuccessful attempts would be audible. The second reason is that this makes the measure more clinically viable. There are problems with this however. Most significantly all non-verbal non-audible interactions are discounted. At times this may make it hard to follow the course of an interaction or a repair sequence. It may even be on occasions that a successful communication involving facial expression gesture or pointing is treated as a failed interaction by the raters as the non-verbal information is not present in the data. Alternatives would be to video the conversations. This would render the measure less clinically useful, and would involve the analysis of verbal and non-verbal data necessitating more time. An interesting future direction might involve the comparison of data from the two recording formats.

These ideas provide some of the directions future research in this area might take, and also provide some clinically useful ideas for dealing with the urgent issue of investigating whether changes in single word production have a correlate in conversation.

11.1.4 Inter-rater and intra-rater reliability

The majority of the variables included in the measure involved a degree of subjectivity. For this reason it was important to establish that two raters coded a particular event in the same category, and that the same rater coded the event in the same category on two separate analyses of the same data. Most studies of conversation data and discourse in conversation attempt to provide such data and typically report the results in terms of proportion agreement. There is no agreed numerical level of acceptability for such data. In the sections below some of the problems with the data and the analysis are outlined and future directions discussed.

11.1.4.1 Level of agreement and the Kappa statistic

Of the analyses carried out here 18 showed acceptable levels of proportional agreement. This is encouraging and demonstrates that in the majority of cases raters shared a common understanding as to how to code the different behaviours. Several showed an unacceptably low proportional level of agreement in either the inter-rater or the intra-rater analysis, and one showed an unacceptably low level of Kappa while having a relatively good proportional level. Raising the cut-off point to, for example 0.70 for the proportional level of agreement, would exclude three more variables (phonological errors, neologisms, and instances of repair). Of these three only one (instances of repair) survived into draft three of the measure, thus in the final version of the measure most variables had agreement levels of over 0.70. In the published literature there is no agreement as to what constitutes an acceptable level. For this measure this is deemed an appropriate level.

11.1.4.2 Sample length

One of the reasons why some of the variables failed to give a satisfactory level of agreement relates to the relatively low number of occurrences of that particular behaviour, which relates directly to sample length. One extreme example of this occurred in repair types, where there were a large number of categories, and relatively few instances of the behaviour. It may well be that with increased sample length, and thus increased numbers of events to code, that reliability of this variable would improve. Again, the question of how long is pertinent. Analysis of one five minute sample using

the original measure was estimated to take two to three hours, which is costly in research time and not feasible in clinical time. One solution, already outlined above (section 11.1.1.2), would be to use a longer sample for tracking occurrence of one particular behaviour known to be of significance for a particular individual. Using a longer sample with the measure described here would be extremely time consuming.

11.1.4.3 Defining terms

The issue of defining what is included and excluded from a category is an important one in this form of data analysis and is crucial if reliability within and across raters is to be established. Many studies fail to provide definitions at all (see e.g. Crockford and Lesser, 1994). In this study definitions of turn types and trouble-indicating behaviours were detailed. Using these definitions substantive turns were agreed well. One reason for this is that the definition was strict: a turn containing a content word was coded as a substantive turn, so, as long as raters agreed upon what was a content word, they would also agree upon these turn types. This definition may however include turns that are not actively contributing new information to the conversation (see discussion of this in 11.1.1 above), and may exclude turns which do contribute new information but fail to contain a content word. An example of the latter would be someone using gesture to communicate new information, which would not be captured by this measure. However turns are defined there will be similar problems in coding agreement, as a degree of subjectivity is always present.

Looking now at topics a less satisfactory picture emerges however. Both inter-rater and intra-rater reliability for topics fell below the cut-off point used in this study. One of the reasons already proposed for this poor level of agreement concerns the identification of sub-topics within a main topic, and whether these are rated as new topics or not. It is possible that defining terms for this variable could lead to better agreement, with all turns that introduce new material not directly related to the previous information under discussion being treated as a topic initiation. Constraining topics within the conversation by requesting speakers to discuss a particular topic known to be of interest to them might also make topic initiation easier to agree upon, as all new information will be within the context of one broader topic. As this is an area which people with aphasia potentially find difficult it is important that some attempt is made to evaluate it.

Certain trouble indicating behaviours which one would have expected to be agreed showed poor reliability and better definitions of terms would have helped here. Inspection of the data showed that uncorrected semantic errors caused disagreements. A definition outlining how to rate such events would have reduced the disagreements here. As most people with aphasia produce semantic errors from time to time it is important that a measure of interactional speech captures this aspect. The case is similar for repetition, where proximity of the repeated element to the initial production could have been defined. Future applications of this form of assessment with people with aphasia would therefore ideally include topic initiation and all forms of error production. In order to reliably code these behaviours stricter definitions of terms would be beneficial.

11.1.4.4 Training

One issue which emerges in the development of an assessment tool is that of training of users. Unless the measure can be used easily by those unacquainted with it, it will be of little benefit except to those who devised it. All three raters who completed the analyses for the reliability study were involved in the development of the measure, thus had developed a familiarity with each other's understanding of the nature of the variables in the measure. It would be interesting to compare a naïve rater's analysis with that of someone experienced with the measure. This is important as generalisability of use is crucial.

11.1.4.5 Summary

The analysis of inter-rater and intra-rater reliability revealed that raters are able to agree well on most variables included in this measure. Some of the aspects of dysphasic speech production (e.g. semantic errors) were difficult to reliably detect however. This is of concern as such a measure of interaction needs to be able to describe the main effects of dysphasic language on the interaction. Aspects of conversation were also hard to reliably detect, such as topic initiation, and repair types. Future work in this area may include stricter definitions, targeted individual analyses over longer conversation samples, and may include aspects of conversation not considered here, such as self-initiated self-repair. These weaknesses notwithstanding, the measure provides a strong first attempt to describe how word-finding deficits affect the conversation of people with aphasia. As such it marks a step-forward in aphasia research and assessment.

11.1.5 Test-retest stability

Test retest stability was analysed by looking at the numerical relationship between scores at two assessment times. As conversation is a relatively unconstrained activity large variations between test times is not unexpected. It may be however that there are some aspects of conversation which remain stable. This is particularly likely with proportional data, as this takes into account the amount of speech a person has produced overall. Thus if someone contributes less to one conversation than to another, within that conversation the proportion of, for example, filled pauses may well be stable. It is therefore important that both raw scores and proportional scores are analysed.

11.1.5.1 Data analysis methods

The sets of scores at the two assessment points were analysed by simple correlations. Although this is standard practice in psychological test development there are problems with this form of analysis. The difficulties are outlined by Bland and Altman (1999), who highlight the fact that a strong correlation can be identified between two sets of scores, even where there are still large differences between the values of the scores. In this case they advocate analysis of the *differences* between the scores, and provide a method of deriving confidence limits for acceptable levels of difference. This form of analysis would be a possible alternative to that used here. The analysis carried out for each individual used the aphasic range of differences and the mean and standard deviation of differences and treated these as the norm. Differences found between scores for the person with aphasia were then compared to these values. This ensured that any change after therapy could be compared to the change found in the baseline scores for the group.

A second possible analysis was used by Nicholas and Brookshire (1993) in their study of aphasic discourse. They used the standard error of measurement to work out how accurately a score at time one predicted a score at time two, and a value 'percentage change' which was derived from the standard error of measurement and gave a measure of the relationship between the two scores. In this study data from twenty people with aphasia were used. This would be a second possible method of analysis for the data collected here.

The use of correlations gives an idea of the relationship between the two sets of scores, but may mask individual differences. For this reason, in considering pre and post therapy scores, the variation in the group of people with aphasia's baseline scores, and in each individual's baseline scores, were considered in relation to any changes found post therapy. This form of analysis ensured that in a case where the group as a whole showed stable scores across the baseline phase, but an individual did not, this latter information was used in the analysis of changes after therapy.

11.1.6 Suitability as a clinical outcome measure

The measure described here represents an important attempt to extend the evaluation of therapy effects from results of assessment tasks into changes in conversation. As conversation is a complex and relatively unconstrained activity, any measure is likely to capture only some of the data. As it is unconstrained it is difficult to show evidence of stability across test times. Although some of the variables in this measure showed group stability across the baseline phase, there were significant numbers of conversation variables which were not stable for a given individual, meaning that any change after therapy was difficult to interpret. Having a larger sample may have been useful, although time restrictions mitigate against this.

One of the reasons why variation is seen across two conversations concerns the topic under discussion. For some topics the person with aphasia may feel able to contribute, whilst for others they may have less interest, be unable to retrieve relevant vocabulary items, or may not have knowledge of the subject. An alternative form of baseline testing might therefore be to take two five-minute samples from one longer conversation. There may still be some problems with topic management, but this would be less likely. The problem with this suggestion is that the next conversation (post therapy) may involve an entirely different topic, which the person with aphasia handles differently to those of the baseline phase.

11.1.7 Future directions

The measure presented here represents a first step towards the development of a valid, reliable and clinically useful assessment tool. Alternatives to the design used here have been outlined. Future work might usefully use longer samples, and fewer variables, and

restrict the conversation to set topics. A means of improving the inter-rater and intra-rater reliability of coding certain aphasic speech behaviours, such as semantic errors, is urgently required. In the meantime, the measure described here can be used clinically, with the proviso that two baseline conversations are recorded, and post therapy changes are viewed in the light of the stability of scores prior to intervention.

11.2 WORD FINDING DIFFICULTIES IN APHASIA

In this study assessment of the language processing and in particular the single word production of three participants with aphasia was conducted in line with a theory of single word processing derived from cognitive neuropsychological studies of similar cases. The theory underlying the methods used is outlined in the model proposed by Patterson and Shewell (1987) and shown in Figure 1.2 in Chapter One. The assessments carried out indicated likely points of breakdown in the processing system for each individual. In this section of the discussion the presentation of each individual is compared to classical theories of aphasia. Points raised by this comparison are then related to theories of spoken word production. Areas which the theories fail to account for are then highlighted. The findings reported here are part of an exploratory investigation of the three participants' word finding. As such no specific hypotheses relating to these findings were outlined at the start of the research.

11.2.1 Classical aphasia syndromes and the data from three people with aphasia

The data collected from PH, SC and KR show a similar percentage of pictures named correctly, and a similar pattern of error types, with high percentages of semantic errors and no responses, and low percentages of phonological errors. Yet the profile of each individual differs from those of the others in significant ways.

Assessment data indicated that PH has a mild semantic deficit, and a deficit in mapping from semantics onto phonological forms. Her sub-lexical routes to spoken output are relatively intact, and act to constrain the nature of errors in the transcoding tasks of reading aloud and repetition, in which she performs well (see Alario et al's 2001 summation hypothesis). She makes mainly no responses, and associative and

circumlocutory errors in picture naming, but also a small number of phonological paraphasias.

SC has a semantic deficit and a probable deficit within the phonological output lexicon. His sub-lexical routes to phonological forms are severely impaired, and do not constrain production in transcoding tasks, thus reading aloud and repetition are errorful. Like PH he makes mainly no responses and associative and circumlocutory errors in picture naming.

KR has a mild semantic deficit, and a deficit in mapping from semantics to phonological forms and is similar to PH in this regard. The deficit in both cases is close to the pattern of pure anomia, although the presence of a mild semantic deficit rules this out⁵³. KR differs from PH however in that her sub-lexical route from input orthography to output phonology is impaired, thus cannot constrain errors in reading aloud. She makes semantic errors in this task. Her sub-lexical route from input phonology to output phonology is relatively intact however. On spoken naming she makes mainly no responses and a range of single item semantic errors (superordinate, co-ordinate, associate, and visually and semantically related errors), thus differs from the two other participants in this respect.

PH's pattern of breakdown fits most closely to descriptions of Benson's (1979) word selection anomia. This form of anomia is associated with the classical syndrome anomic aphasia, where the predominant symptom of aphasia is the marked word-finding problem. Speech is fluent and comprehension is spared. Speakers make semantic errors but no phonological errors, and they can describe items they are unable to name. PH's profile is similar to this apart from her mild semantic comprehension deficit, and the fact that she does make some phonological errors.

In semantic or nominal anomia (Benson, 1979) comprehension of single words is impaired through both the written and spoken modalities, speech is fluent, and there are many semantic paraphasias and paragrammatisms. This is associated with Wernicke's aphasia. This pattern fits the presentation of SC, although his semantic comprehension

⁵³ The evidence for a semantic deficit comes, especially for PH, from the effects of semantic variables on naming. Analysis of the effects of variables is not routinely carried out in all studies of aphasic naming.

deficit is not severe. He makes semantic errors, and he produces paragrammatisms in everyday speech.

Benson (1979) described word production anomia occurring in the context of Broca's aphasia, with effortful non-fluent speech, and relatively good comprehension. Errors in single word production are commonly phonological distortions. Goodglass (1997) reports an over-reliance on nouns in connected speech. This description is the closest approximation to KR's pattern of speech production. Analysis of her noun production in conversation shows that she produces proportionally more nouns than do the other two participants (see Tables 10.12, 10.13 and 10.14). She differs from this classical account however in that she has a mild comprehension deficit, shown in tests of semantic processing, and she produces mainly semantic errors and not phonological distortions.

Although according to Benson's (1979) criteria PH and KR fall into different categories, it is worth noting that purely in terms of their word finding difficulty they are similar. This point highlights the weakness of the syndrome approach, and underlines the need for more detailed assessment such as was carried out here.

11.2.2 Theories of spoken word production and the data from three participants with aphasia

The theories of production outlined in Chapter Two need to be able to account for the data from people with aphasia. One restriction to the application of these theories to understanding aphasic word finding lies in the uni-modality account they provide. None of the models outlined in Chapter Two (Caramazza, 1997; Dell et al, 1997; Levelt et al, 1999; Rapp and Goldrick, 2000) provide any account of input processing, of routes bypassing semantics, or sub-lexical routines. As a result they are unable to explain in their entirety the patterns of single word and non-word processing found in the participants with aphasia described here. They can only account for patterns of spoken word production in isolation from all other data. Nevertheless they should be able to describe the patterns found in picture naming in these three participants.

11.2.2.1 Levels of processing

In this section Dell et al's (1997) DSMSG model and Levelt et al's (1999) WEAVER++ model, which both incorporate a syntactic level mediating between semantics and phonology, are compared with Caramazza's (1997) Independent Network Model, in which access to phonological lexemes proceeds directly from semantics. All three models can be damaged at the level of lexical semantics. Although the architecture of the lexical semantic level differs across the models, with distributed semantics in some theories and non-decompositional semantics in others, damage to this level would, one assumes, lead to mis-selection of a word or a lemma in two-step models, or of a phonological form in one step models. It is logical to assume that the mis-selected item would have some semantic relationship with the intended target.

Explanations of difficulties in mapping from semantics to output phonology differ however across the models. In the data collected for this thesis there are two participants with a similar pattern of performance on tests of language processing, but who make different types of semantic errors (PH and KR). A single step model (e.g. Caramazza's Interactive Network, 1997) which maps from semantics straight onto phonology would, presumably, result in the same types of errors for all people with aphasia with damage in the mapping processes. Two participants with similar sources of damage but different patterns of errors are problematic for such a model. The presence of an intermediate syntactic stage allows a further differentiation between these aphasic speakers, with some having problems accessing the word or lemma level, some having a deficit within the word or lemma level, and others having a deficit in mapping from the word or lemma to the phonological level. This allows three possible patterns of deficit where a one step model allows only one possibility. It is not clear how the three different forms of deficit in a two step model would present in terms of error patterns but some tentative conclusions can be drawn from the error data of PH and KR. If a participant were accessing the syntactic level effectively errors would be constrained syntactically, that is, if a noun were the intended target the error would be a noun (Dell et al, 1997). The predominance of circumlocutory and associative errors found in PH's naming, many of which productions included verbs and adjectives, suggest that there is no syntactic constraint upon her responses, which in turn indicates that she is not routinely accessing

the syntactic level, or that there is damage within that level⁵⁴. KR on the other hand produces mainly other nouns, which are related by meaning to the target. This suggests that she is accessing the syntactic level and that this is constraining the types of errors she can produce.

SC has damage at the lexical semantic level, and within the phonological level. Like PH however, his semantic errors are also mainly associative errors and circumlocutions. The lack of a syntactic constraint upon his error production suggests that he also has damage at the syntactic level. This would necessitate hypothesising damage at all three levels in order to explain his spoken word production. The lack of phonological errors in his naming is problematic here, given that he has putative damage to output phonology. Information from his performance in Phase One therapy is pertinent here however. Recall that in the picture naming assessment performance was scored by taking the last response within the first five seconds of seeing the picture. So the data from all of the participants reviewed here refers only to what they produced in those first five seconds. In Phase One therapy participants made multiple attempts at a target as the cueing procedure progressed. In this scenario, with the help of first sound or first letter cues SC produced many non-word phonological approximations to the target. This provides further support for a phonological level deficit, and highlights the fact that the criteria used for assessment can lead to different readings of the data⁵⁵.

The issue of what role the intermediate syntactic level plays in a task requiring production of nouns in isolation remains open to question. It may be that access to syntax facilitates word production. In the absence of this syntactic activation naming is

⁵⁴ An alternative explanation for PH's errors would propose that her single word verbs and adjectives constitute attempts at a circumlocution. Counter to this claim is the fact that PH has fluent spoken output and therefore is capable of more than a single word circumlocution.

⁵⁵ A different explanation for SC's output phonology might therefore be necessary. Recall that the evidence for SC's phonological output damage comes from his poor performance on tests of reading aloud and repetition. An alternative hypothesis to that of phonological output damage might be as follows: in tasks without obligatory semantic involvement, such as reading aloud and repetition, there is insufficient semantic drive to maintain the integrity of phonological representations. This would mean that these are intact and available *but only* when sufficient semantic activation is present. In picture naming on the other hand activation from semantics leads to the production of the target or of a semantically related word, the phonology being activated by the semantic activation. This explanation has been used to explain the gradual breakdown in phonological output found in cases of semantic dementia (e.g. Hodges et al, 1992). If one accepts this account SC's pattern can then be explained as a more severe type of the deficit seen in PH. Both have some semantic impairment, both produce similar types of semantic errors, but for SC there is insufficient semantic activation to allow good performance on tasks not overtly requiring semantic involvement.

further impaired than it would otherwise have been, and there is no constraint acting over the syntactic class of errors, with associative errors and circumlocutions appearing.

11.2.2.2 Activation

The theories also differ as to how activation spreads through the system. Discrete feed-forward activation found in Levelt et al's (1999) theory predicts that there is no influence of subsequent levels of processing on previous levels. In interactive activation models there is an influence of later levels on processing at previous levels. Interactive activation was built into Dell's (1986) original model in order to explain two features of normal speakers' speech errors: the mixed error effect, and lexical bias (see Chapter Two, section 2.3.7.2). The two forms of activation provide different predictions for the patterns of damage hypothesised to be present in PH, SC and KR. In the discussion offered below damage is viewed as having a catastrophic effect on the system. This is assumed in order to be able to compare each person's profile to the predictions the models offer. In reality the three people with aphasia present with graded degrees of damage.

If PH has damage to semantics, and to the syntactic level of processing, a discrete feed forward model would predict that she would make semantic errors, that these would not be constrained by syntactic class, and that she would not make phonologically related errors. This is because once selection or mis-selection at the syntactic level was completed, processing would continue forwards, to selection of a non-damaged phonological form. There would also be no evidence of mixed errors (errors related in meaning and form to the target), and no production of non-words. An interactive activation account would predict semantic errors, again not constrained by syntactic class, but that feedback from the phonological level to the syntactic level would lead to higher activation of items related in form to the target. Thus formal errors and mixed errors should occur. Non-words should not occur as processing between the syntactic level and the phonological level is intact. As noted above, PH does make semantic errors and these tend to differ in syntactical class from the target. Both forms of activation cope with this. PH does however make some phonologically related errors, which only the IA account predicts. She does not however make mixed errors, contrary to the prediction of this account: in each assessment she made only one error which was both semantically and phonologically related to the target. Thus PH's profile is

adequately explained by both accounts but each has its area of difficulty. The feed-forward account does not predict PH's production of phonologically related errors, and the IA account wrongly predicts mixed errors.

SC has a profile with hypothesised damage at all three levels of processing. A discrete feed-forward account predicts semantic errors, no syntactic constraint upon semantic errors, phonologically related errors with no syntactic constraint, and non-words (due to phonological damage). The phonologically related errors should also show no evidence of a syntactic constraint as there is no interaction between the phonological level and the syntactic level. The IA account predicts semantic errors, no syntactic constraint upon semantic errors, phonologically related errors, and, because of damage to both the syntactic level and the phonological level, no syntactic constraint on phonologically related errors, and non-words. The latter occur because the interaction between syntactic and phonological levels is impaired due to damage at both these levels, thus words do not win out over other phonological strings. Both accounts predict the same results therefore. This suggests that in cases where there is damage throughout the system, and where the interactive mechanism cannot function effectively, there is no difference between the two accounts.

KR has mild damage to semantics and further damage to the processes mapping from an intact syntactic level to phonological forms. The discrete feed-forward account predicts that semantic errors would occur, and that they would be within grammatical class. In addition, due to damage to processes mapping between syntax and form, formal errors should occur. This is because the selected lemma or syntactic node would not map onto its corresponding phonology. A close neighbour might then be activated. This neighbour need have no semantic or syntactic relationship to the target, as processing at these levels is completed before the activation of the phonological form commences, but such a relationship might exist by chance. Non-words should not occur as there is no damage to output phonological representations.

For KR the interactive activation account would predict semantic errors, again within grammatical class, as activation between these two levels would constrain selection at the syntactic level. Damage to interactive processes between syntax and phonology would lead to the activation of formally related words at the word level, leading to formal errors. Non-words would occur as there is a reduction in the degree to which the

syntactic level and phonological levels interact, thus, the usual advantage for strings of phonemes which are activated by a node at the word level, over strings of phonemes lacking such activation, i.e. non-word strings, would not be present. In the normal system formal errors should obey the syntactic class constraint. In a system where there is damage to this interaction this constraint would not necessarily operate.

KR produces semantic errors and these are constrained by syntactic class. She produces few formal errors, and as the number of these is so small it is not possible to gauge whether there is a syntactic class constraint operating over their production. Where the two sets of predictions differ is in the production of non-words. The feed forward account predicts no non-words, whereas the IA account predicts that there will be non-words. KR produces very few non-words and thus the prediction of the IA account is not supported.

One caveat to the above proposals concerning PH and KR concerns the overall language profile of each. It is proposed here that PH's deficit lies in access to or within the syntactic level, and KR's in accessing phonology from an intact syntactic level. PH however has fluent spoken output which contains syntactic structure including function words and grammatical affixes. KR's spoken output is agrammatic with a lack of syntactic structure and grammatical affixes. The proposals outlined here appear contradictory in this larger framework. The account proposed here is however restricted to noun production. It is feasible that noun syntax could be disturbed in a person with fluent aphasia, and not so in someone with non-fluent aphasia. The disturbances seen in sentence production may relate to processing at a different level/s e.g. mapping from sentence semantics to syntax, or be due to a deficit in verb processing (see e.g. Berndt, 1998 for a discussion of aphasic sentence processing).

11.2.2.3 Summary

In this section the ability of three models to account for the data from three people with aphasia is discussed. In terms of levels of processing it is clear that a two-step account with three levels of processing offers more possible loci of deficits and is more able to account for the variation seen across participants with aphasia. In particular where two people present with the same putative locus or loci of damage according to the one-step model, but their error patterns differ, the two-step account can explain the two profiles

more readily. Crucially the two-step accounts incorporate an intermediate stage at which syntactic information is accessed. Dell et al (1997) provide a detailed account of the impact of this level of processing upon adjacent levels in normal speech production. They do not however detail how damage to this level or to the processes linking it to other levels will manifest itself in aphasic speech production. In their simulation damage to the system is global, and therefore discrete levels of processing are not implicated. There are a small number of studies which have looked at syntactic relationships between semantic errors and targets in aphasia. Kulke and Blanken (2001) found that aphasic semantic errors preserved the target's grammatical gender at rates above chance. The data came from picture naming where single word production only is required and points to the possibility that even when syntactic information is not overtly required syntax is nevertheless activated. This area urgently needs single case study investigations of levels of processing of spoken word production in aphasia, including the syntactic level, in order to begin to build up an account of likely processing breakdowns involving that level.

In Dell et al's (1997) simulation of aphasic naming the naming profiles of participants with aphasia were reproduced by the computer model by damaging two of the model's parameters: connection weight and decay rate. Each person with aphasia could then be explained by the model in terms of either a reduction to the former or an increase in the latter, or both. All of the participants produced no responses at rates of less than 15%, and no responses were not included as an error category. This means that the data from the three participants with aphasia who took part in the study reported in this thesis cannot be considered within this model, as all produced no responses at rates well above 15%. Dell et al (2004) have produced a new version of the model which incorporates non naming responses⁵⁶. In this category they include semantic circumlocutions and visual errors. The form of the model which is best able to fit the data from people with aphasia who produce non naming responses is the semantic-phonological model (Foygel and Dell, 2000) with either an assumption that non naming is independent of the overall processing system, or that non naming arises when no lexical item reaches threshold. These variations of the semantic-phonological model are more able to accurately reflect the range of aphasic performance, but the assumption that non

⁵⁶ The model is based on the naming performance of normal participants and people with aphasia on the Philadelphia Naming Test. Performance on other tests of naming is not directly comparable and therefore it is still not possible to compare the participants who took part in this study with those described by Dell et al (2004).

naming, semantic circumlocutions, and visual errors emerge from the same processing breakdown is open to question.

11.3 THERAPY

In the study described here the three participants underwent two forms of therapy: a lexical therapy using phonological and orthographic cues (Phase One), and an interactive therapy using a set of treated words in quasi-real speech situations (Phase Two). In the first part of the discussion the effects of the lexical therapy will be discussed against the background of the language profile of each participant. This will be followed by a discussion of recovery in terms of Robertson and Murre's (1999) criteria for effective rehabilitation. In the second part of the therapy discussion the effectiveness of Phase Two therapy is discussed. Finally the effects of therapy on conversation are described.

11.3.1 Phase One Lexical therapy

11.3.1.1 Relationship between therapy and level of impairment

Of overwhelming interest in the field of aphasia therapy is the quest to construct a theory of therapy. Such a theory would be able to identify for any given individual therapies likely to be effective for them. This consists of comparing the language profile to the therapy effects in case series studies (e.g. Nettleton and Lesser, 1991; Hickin et al, 2002). One issue of concern in this regard is which aspects of language processing need to be taken into account in the description of any given individual. Here theory may help. In looking at spoken word production certain elements appear more relevant than others: semantic processing is assumed to be more relevant than, for example, digit span, or syntactic production. Thus there is a strong selection bias present in any study. In the work reported here five elements of processing are considered with regard to therapy outcomes: semantic processing, mapping from semantics to phonology, phonological representations, sub-lexical processes (phonological and orthographic routes to phonological output), and the response to cues. In the previous section participants' performance was compared to a model incorporating an intermediate level between semantics and phonology. No specific tests of processing at this level were

conducted. In the account that follows this aspect of processing will be considered theoretically.

The specific hypotheses relating to Phase One therapy proposed that this form of therapy would be most effective for participants with good semantic and phonological processing, and a deficit in mapping between the two levels. It was further hypothesised that as the therapy operated on the mapping between the two levels, item-specific effects of therapy would be found.

11.3.1.1.1 PH Phase One

PH has a mild semantic deficit, a difficulty mapping between semantics and phonological output representations, possibly involving syntactic representations, and she responded best of all three of the participants to cues. A reasonable prediction (see above) might be that the Phase One therapy would be effective for her, and so it proved. She made significant gains in picture naming, and both phonological and orthographic cues were effective. There was no effect on untreated items. The cues as they appear are sub-lexical in nature, therefore participants' sub-lexical processing is possibly of interest here. PH had relatively good sub-lexical processing (see Table 10.1). If her main deficit lies in achieving the mappings between semantics and phonological output representations, it may be assumed that the therapy strengthened these links, using her strong ability to produce lexical items given a sub-lexical input stimulus (which is what the phonological and orthographic cues amount to). It is interesting that PH responded so well to this form of therapy given that, as hypothesised in section 11.2.2, the most parsimonious account of her spoken word production deficit postulates a deficit in or in accessing the syntactic level. Just how this form of therapy affects processing at this level is unclear. Indeed the results of the therapy support the initial proposal (Chapter Seven, section 7.10) of a main deficit in accessing phonological representations from semantics. Of the reported cases of phonological therapy PH is most similar to RBO (Miceli et al, 1996) who also had a severe word finding deficit, and whose main deficit lay in mapping from semantics to phonology. Miceli et al (1996) predicted that the phonological therapy would restore access to unavailable output lexicon representations and, as therapy focuses specifically on particular items only, no effect will be seen in untreated items. This was the case for RBO as it was for PH. Just how important sub-

lexical processing is for this form of therapy to work will be discussed further in relation to SC and KR's profiles.

In summary the effect on PH's word finding was in line with the overall hypotheses concerning level of breakdown and response to therapy, and effects of therapy on treated and untreated items. The therapy had a significant effect on picture naming; only items treated in therapy gained.

11.3.1.1.2 SC Phase One

SC has both a semantic deficit and a phonological deficit in spoken word production. He may also have a deficit at the level of syntax. In addition he showed no response to cues. A reasonable prediction here would be that the lexical form of therapy would be ineffective. Again, this proved to be the case. Neither treated nor untreated items showed any improvement. There was no effect of either form of intervention on his word-finding. SC's phonological output representations were hypothesised to be damaged. This may mean that any form of therapy which seeks to improve access is not going to be effective for him. An alternative form of therapy which seeks to work on the representations themselves, such as that used with GF by Robson et al (1998), may be more beneficial. In this therapy GF was asked to think about the syllabic structure and first phoneme. A further factor in SC's lack of response to the therapy is his sub-lexical processing, where he showed a marked impairment. He did poorly in non-word repetition, and his orthographic to phonological conversion route showed no retained function at all. Just how important a factor this level of processing is in this form of therapy is unclear. It would be of interest to compare someone with a mapping deficit similar to PH's, which responds well to this form of therapy, but who also has a sub-lexical processing impairment. This was partially the case with KR.

As for PH the results for SC are in line with the original hypotheses. SC's deficit in spoken word retrieval involves damaged semantic and phonological processing, and the therapy was not effective for him. A therapy which acts by improving links, or mappings, between adjacent levels, cannot also remediate impaired processing within those levels.

11.3.1.1.3 KR Phase One

KR has a mild semantic deficit, and a difficulty mapping between semantics and phonological output representations, with retained access to the syntactic level. Unlike PH but like SC she demonstrated no response to cues. She showed good repetition and thus it may be assumed that phonological output representations are intact and accessible. It is harder to predict how KR would respond to the lexical therapy given the information above. In the event she made significant gains in picture naming, although not to the same extent as PH. Like PH the treatment effect was confined to treated items. This suggests that the same mechanism was operating and in Miceli et al's (1996) terms this implicates the links between specific semantic representations and their corresponding phonological output representations (they consider only a one-step model). So far so good. The story grows even more interesting however when the differential effects of the two forms of cues are considered. For KR there was a significant effect of the phonological cues but a smaller effect of orthographic cues in therapy. Corresponding to this pattern, KR has some retained function in sub-lexical conversion of input sound to output phonology. Like SC she has no retained function in sub-lexical grapheme to phoneme conversion: she was unable to produce any phonological output at all in this task. Thus the cue type which had an effect is supported (hypothetically) by retained function in that domain. The cue type which was ineffective is unsupported, with complete obliteration of that process. Why KR's performance with orthographic cues improved at all remains a mystery therefore. The small improvement could be due to random variation in performance, or could result from KR covertly extending her use of the phonological cues to the orthographic set⁵⁷.

The findings from KR are therefore also in agreement with the hypotheses concerning Phase One therapy. KR has a deficit in mapping between relatively good semantic and phonological levels, and the therapy was effective for her. In addition the therapy effect was restricted to treated items. An additional finding which was not predicted concerns the relationship between the response to therapy cue types (phonological and orthographic) and processing in the corresponding sub-lexical routines.

⁵⁷ Against this hypothesis is the evidence that for KR there was no effect on untreated items,

11.3.1.1.4 Summary

The components of language processing which appear to be crucial for an effective response to this form of therapy are: fairly intact semantics, intact phonological output representations, and some function in the relevant sub-lexical processing routines⁵⁸. As no assessment of the syntactic level was conducted it is not possible to hypothesise further about this level of processing: evidence for the retention of processing at this level has come entirely from whether a syntactic constraint operates over error production (see section 11.2.2) and is therefore not robust.

Just what level of semantic involvement is significant is open to debate: both PH and KR have some impairment at this level. It may be that a combination of factors is key. If someone presents with *some* semantic impairment, but has access to phonological output representations in, for example, word repetition, and can convert input stimuli (either letters or words) to output phonology at a sub-lexical level, then the therapy may well be effective. This hypothesis challenges the notion that therapy should target the functional level of damage. It suggests that a combination of factors is of more importance. What these factors are for any given therapy remains to be determined but those identified here provide a start for phonological/orthographic therapies. In most reported cases in the literature those people with aphasia undergoing phonological therapy are people with little semantic damage, thus there is no opportunity to investigate the level to which semantic processing needs to be preserved. In future reports combinations of factors need to be analysed and alternative forms of statistical analysis used which can accommodate multiple factors.

11.3.1.2 Components of therapy

All three of the people described in this thesis were many years post onset of their stroke and therefore no spontaneous recovery was likely. Just how recovery of function

⁵⁸ In the related published report of the lexical therapy (Hickin et al, 2002) where the results of the therapy for eight people with aphasia, including the three described here, are described, significant correlations between improvement in naming all the 200 items and the following factors were found: written word to picture matching (a semantic factor) and initial phoneme production in reading non-words (a sub-lexical factor). In addition a significant correlation between improvement in the phonological set and performance in written word to picture matching was found. Although not supporting the hypothesis outlined here entirely, in particular as no relationship between improvement in picture naming and scores in tests of repetition was found, these results show partial corroboration of the relationship between response to therapy and intact semantic and sublexical processing.

operates in such a scenario is of interest. Robertson and Murre (1999) propose that maximal recovery will depend upon i) principled stimulation based on theoretical models of mental processing, ii) significant and adequate amounts of repeated administration of the stimulation, iii) awareness of deficits and the attention of the person being directed to the stimuli, and iv) adequate arousal levels in the person concerned. Without all four of these criteria guided recovery will not be effective.

11.3.1.2.1 Principled stimulation

In the lexical therapy conducted here it was hypothesised that access to the (intact) phonological output level representation could be improved by targeting the first sounds and letters of the word. This assumed that there was adequate semantic activation and that the corresponding phonological output representation was available. Thus the therapy was based upon a simple functional model of access from semantics to phonology in one step. As more complex models emerge, such as the two step model of Levelt et al (1999) and Dell et al (1997), and as tests of the validity of models with alternative forms of activation to simple feed-forward processing are presented (e.g. Rapp and Goldrick, 2000), alternative forms of therapy should also emerge. Support for the notion that the therapy was based upon an adequate model comes from the fact that the therapy predictions based on the participants' language profiles were on the whole correct.

11.3.1.2.2 Amount of stimulation

The therapy consisted of one session per week for eight weeks. Each item was therefore presented for naming eight times only. For PH this was an adequate amount of therapy but it is arguable whether it was sufficient for KR, the other participant who showed some response to therapy. Alternative designs could have increased the number of sessions per week, making the therapy more intensive, the number of sessions overall making the therapy of longer duration, and the number of times an item was presented in each session, thereby increasing the 'dosage'. In the therapy described here it was possible to track the overall number named correctly, and the number of items responding to the different cue levels, in each session. Thus in a clinical setting it would be feasible to increase the intensity or the amount of exposures of each item as appropriate dependent upon scores in each session. What is not in doubt is that repeated

stimulation of a relatively small set of items is crucial. Using different items in each session will not achieve the same results: untreated items did not gain, and the therapy effect overall was gradual across sessions. It is also of note that people with aphasia find repetition of the same set of items challenging. SC was aware of which words he could and could not get and was keen on trying them repeatedly. After one session he said: “As you’re constantly doing the same one each time, I’m doing them more of. All of a sudden I might have succeeded for the first time”. KR felt there were enough words to work on and that doing the same set repeatedly was beneficial: “Every day the same thing over and over. he help me”.

11.3.1.2.3 Awareness of deficits

The three participants who are described here all had good awareness of their word-finding deficits. Robertson and Murre (1999) report this factor to be critical as without this awareness “attention will not be directed to inputs that might play a role in facilitating plastic reorganization of the brain” (Robertson and Murre, 1999: 563-4). They claim that plastic neural reorganization does not occur passively but requires the active attention of the participant to the stimulus. The therapy conducted here conducted overtly in that the participants were aware of the aims of the intervention and their attention was drawn to the stimuli. As a result their attention was directed to the process of spoken word production and to the cues as they were presented. In many cases of aphasia this awareness and consequently this level of attention are not present and in such cases rehabilitation efforts need to be devised in line with this. According to the theory someone with a poor awareness of the deficit and poor attention to the stimuli should not improve in therapy.

11.3.1.2.4 Arousal levels

Overall arousal levels are also deemed to be a critical factor. Without active attention being paid to the stimuli successful rehabilitation will not occur. In order to attend a significant arousal level needs to be sustained throughout the therapy session. This point recalls that made above regarding the amount of therapy offered in each session, as long sessions may be inappropriate for certain people, and short and often may give them a better chance of sustaining attention throughout. In a therapy such as the lexical therapy described here it is therefore crucial that the therapist is alert to the possibility of a drop-

off in performance, and checks frequently with the participant regarding their engagement in the task. This is not performed routinely but is strongly advised.

11.3.2 Phase Two Interactive therapy

In Phase Two a set of items were treated in situations which gradually approached more natural speech production. There are relatively few reports of similar therapies. Springer et al (1991) described what they termed a modified PACE therapy. This involved a semantic categorisation task combined with the traditional PACE task of communicating information to a second party who is ignorant of the content of the message to be conveyed. Their results showed a superior effect of the modified PACE task over a traditional PACE task (where the person merely communicated the sense of the picture to their therapist). In Springer et al's study the lexical semantic component of the modified task distinguished the two approaches. In the Phase Two therapy reported here participants were able to use cues as they wished to facilitate word production but this was not obligatory. It is not clear which aspects of language function might predict the response to this form of therapy, as it involves selection of lexical items to convey a message, thereby calling upon multiple linguistic and conversational skills. The overall gains were measured in terms of picture naming performance however.

Specific hypotheses relating to Phase Two were that the therapy would be effective for a broader range of deficits in word retrieval than was the lexical therapy; and that generalisation to untreated items would occur.

11.3.2.1 Overall effectiveness of the therapy

As this therapy does not target individual lexical representations, but rather the production of lexical items in interaction, it might be predicted that the therapy effect would generalise to untreated items and to conversation. The first proposition was not upheld. The therapy effect was, like that of Phase One, restricted to those items appearing in therapy. This suggests that, although the intention of the researcher was to encourage communicative use, the actual effect of the therapy was similar to that found in Phase One, with specific words receiving stimulation and thereby becoming more accessible.

11.3.2.2 Individual responses

More encouragingly than Phase One, all three participants responded to this form of therapy showing gains in picture naming of treated items. This is in line with the hypothesis that the therapy was operating on multiple levels of linguistic processing, and thus would be effective for a broader range of deficits. For PH and KR it may be that the same mechanism that led to gains in Phase One therapy also led to gains in Phase Two. They both have a deficit in mapping from semantic to phonological representations (but see comments above in section 11.2.2 and 11.3.1 regarding the involvement of the syntactic level) and it may be the case that both forms of therapy facilitated this mapping, albeit in different ways. For SC this therapy response is intriguing. He showed no response at all to Phase One therapy, and although he did not make gains in naming overall, his naming of treated items improved significantly after Phase Two. This was offset by a deterioration in performance in one of the untreated sets.

For SC it is interesting to explore why Phase Two was effective. In those items which received stimulation in Phase Two there were significant changes. In a clinical setting items of real worth to SC in everyday life could thus be usefully targeted using this form of therapy. One difference between Phase One and Phase Two is that for Phase Two items were selected as far as was possible according to their functional usefulness to the person. This was not the case for Phase One. Thus the items treated in Phase Two were those which SC would be more likely to need to use in everyday life. It may be the case that for SC the activation levels of words are so suppressed that no amount of lexical therapy will be effective. Of interest would be a follow-up comparison to investigate whether the set treated in Phase Two only would respond to Phase One therapy. SC had a deficit within phonological representations. In order to ensure that items treated in therapy had a fair chance of improving it would be worth investigating word repetition of treatment items prior to embarking upon therapy to ensure that access to the phonological representation was possible at least under certain conditions.

11.3.2.3 Response of the sets to therapy

Of note in this phase of therapy is the fact that for all three participants the most marked improvement occurred in those items which had not been treated before. For both SC

and KR the only significant improvement occurred in this set. For PH this set improved and the set treated in both phases also improved but to a lesser degree. It is possible that a ceiling effect is operating here. Items treated in Phase One had reached their ceiling beyond which they could not improve. It is possible that within a given set of randomly selected words, an individual with aphasia can reach their own maximum score beyond which they cannot go. This is determined by the nature of the set and the psycholinguistic variables which affect that person's word finding. This score will vary across individuals naturally, and within a given individual, across sets (as these differ in their make-up).

For the three participants there was a possible ceiling also in the overall number of items named correctly out of 200. In all cases improvement in one set was offset by deterioration in another set. The majority of this change was found in the untreated sets, the exception being KR who also showed some deterioration in the set treated in both phases (0.50 to 0.46). As therapy acts on one subset of the items it pushes certain items to a state where they are more readily accessible. For any given individual the items which respond to therapy will in all likelihood be governed by the same variables which affect naming overall. Once the maximum score is reached however, gains in the therapy set will be offset by losses elsewhere (to balance the books). This implies that in selecting items for therapy it is essential that due consideration is given to variables affecting a person's word finding. What is not clear from the present state of knowledge is which items will be more likely to respond. If someone's word finding is affected by frequency for example, it is not clear whether therapy should target high or low frequency items preferentially.

The second possible explanation for the deterioration in untreated sets relates to the process of allocating items to sets in Phase Two. Items were matched for baseline and post Phase One therapy performance, but items which were deemed functionally relevant were selected for the treated sets (tT and uT). By definition the items in the untreated sets (tU and uU) were therefore less functionally relevant. In selecting the items many everyday objects, foods and so on were considered functionally relevant, thus the treated sets may well have been of higher familiarity and higher word frequency than the untreated items. This was investigated for SC and found not to be the case, however it may have been so for the other two participants. This does not negate

the value of this form of therapy, as successful use of everyday terms is a highly desirable outcome in itself.

11.3.2.4 Distinguishing between Phase One and Phase Two effects

The design of this study in which two therapies are conducted consecutively does not allow one to distinguish the independent effects of Phase Two therapy. All the participants participated in Phase One and this may have affected their response to Phase Two. For example it is possible that SC's good response to Phase Two resulted from a delayed response to Phase One. As only those items which were untreated in Phase One and treated in Phase Two improved, this is however unlikely. Alternative designs would present the two therapies in parallel, or would vary the order of presentation of the two therapies across participants. This is counter-intuitive however, as the lexical therapy should ideally serve as a means of setting up the better production of lexical items, and the interactive therapy should then work on the production of those items in everyday speech.

11.3.2.5 Maintenance of therapy effects

The design of this study allowed maintenance of therapy effects to be measured at a period of two months after the completion of Phase Two therapy. This design meant that overall maintenance in the set of 200 pictures, and effects in control tasks and other measures, notably conversation, could only identify the cumulative effects of both phases of therapy. In order to track maintenance of Phase One therapy in isolation the set of items treated only in Phase One was analysed. Again there is a problem with claiming that effects seen in this set are entirely due to Phase One therapy, as generalization of therapy effects from Phase Two may well have affected production of items in this set. Against this hypothesis is the evidence that the predominant therapy effect was found in items treated in Phase Two only.

Viewing the set of 200 items as a whole, all three participants maintained their performance from assessment four, to assessment five. In PH and KR's case the significant changes overall across the two therapy phases were maintained. In SC's case there was no change after either phase of therapy, thus overall performance merely stayed stable throughout SC's involvement in the study.

Viewing the set of items treated only in Phase One (tU), for PH there was a gradual deterioration in this set after Phase One therapy. There is little evidence here therefore for claiming that the specific effect of Phase One therapy lasted. The time lag between assessment three and assessment five is approximately four months. It may well be that items treated in therapy need to be practised consistently in order for improvements to maintain. Once therapy is terminated activation levels which had been raised by therapy (according to, for example, Dell et al's 1997 account of aphasic naming) return gradually to their original state. This is disappointing, particularly as there is evidence from other studies, notably from Miceli et al (1996), showing lasting effects of phonological therapy.

SC showed no improvement in the tU set after Phase One and he performed at the same level on that set throughout. KR on the other hand showed significant improvement in the tU set after Phase One therapy. Her performance on this set was erratic thereafter however. At assessment four she dropped back to her pre-therapy level of performance. At assessment five she recovered to the same level as assessment three. These data are hard to interpret. There is no convincing case for claiming that KR's performance maintained at follow-up. There is variation in her data which suggests that multiple assessment points are required to identify trends within a series. Inconsistency in individual people may need to be identified before the design of a therapy study is determined. In KR's case this would argue for multiple baselines greater than two assessment points. For other participants this may not be necessary.

In conclusion there is some evidence in two cases of maintenance of improvement in the set of items overall. There is little convincing evidence in the analysis of the individual sets undergoing the two therapies of maintained performance. One reason why this is hard to tease out relates to the design of the study, in which the effects of the two therapies are confounded. It would be worth carrying out an investigation of each of the two therapies in isolation from the other in order to untangle these effects.

11.3.3 Conversation

In this final section the degree to which it is possible to track therapy gains in terms of changes to conversation is discussed. The design of the conversation measure and

possible adaptations to that measure have been outlined above. Here the individual scores in the conversation measure after therapy are discussed in the context of the original hypotheses concerning the relationship between the two forms of therapy and likely changes in conversation. It was hypothesised that there would be no effect on conversation as a result of Phase One therapy, but that carryover to conversation would occur after Phase Two therapy.

11.3.3.1 Number of variables and effects of therapy

In the third draft of the measure described here there were 12 variables based on raw scores and six variables using proportional data. Given these 18 sets of figures it is possible that an apparently real change could present which is in fact a random presentation (type one error). The data need to be viewed with extreme caution. In addition the measure represents a first attempt to capture changes found in picture naming of 200 items in conversation, a conversation which may never require the production of any one of those 200 items. The measure is thus a fairly insensitive tool. A related point concerns the fact that both therapies for all three participants resulted in improvement (where this occurred) in treated items only. There was little evidence of improvement in untreated items. This lack of generalisation is important here as, if generalisation were present, this might result in changes in conversation. As it is no such change was found, thus the conversation data was trying to capture an improvement in only 50 or 100 words. This is unlikely to happen.

11.3.3.2 Effects of therapy on conversation

For two of the participants (SC and KR) there were unstable baselines for the majority of the conversation variables. This means that it is very difficult to interpret the post therapy data. For PH however, the majority of the variables were stable prior to therapy, and thus it is possible to view changes after therapy with some degree of confidence. Contrary to the original hypothesis, which predicted no effect of Phase One therapy on conversation PH showed an increase in lexical retrieval and substantive turn production. After Phase Two there were some positive changes e.g. fewer filled pauses, but some negative changes as well, e.g. more minimal turns, and fewer nouns. There was thus no overall positive effect of Phase Two therapy on PH's conversation. PH showed the strongest response to Phase One therapy and it is encouraging that there is

some evidence of this affecting conversation. Less encouraging is the evidence that after Phase Two therapy, which was hypothesised to lead to improvements in conversation, there is little consistent evidence of this occurring.

A possible alternative to the design used here is to track changes in one particular behaviour over a number of conversations. In PH's case it would be feasible to investigate substantive turn production using multiple baselines combined with multiple post therapy assessments. Thus the overall score and the degree of change in the baseline phase could be compared to the overall score and degree of change in the post therapy phase.

11.3.3.3 Summary

The conversation measure was used to directly assess the effects of therapy on specific aspects of conversation relating to word-finding. In most studies of therapy for word-finding outcomes in terms of picture naming are reported but there is no attempt to measure everyday use of language. In some studies connected speech tasks have been used (see Chapter Four, section 4.2.3.1.4) but most of these lack ecological validity. The measurement of conversation proved to be problematical. Most importantly test-retest stability was not present for a majority of variables for two of the participants. For one participant the majority of variables showed adequate test-retest stability to allow tentative conclusions to be drawn about the effects of therapy on conversation. This participant's picture naming also showed the greatest therapy effects. The evidence of positive changes in conversation after Phase One therapy is compelling. It may be that in order for such changes to be identified a large therapy effect (in terms of gains in picture naming) needs to be present. Nevertheless these results are extremely encouraging.

11.4 MAIN FINDINGS OF THE RESEARCH

Lexical therapy, which focuses on word form (phonological or orthographic), is an effective means of improving word finding for at least some people with aphasia. The effect of this form of therapy is item-specific with items not undergoing therapy showing no improvement. This was predicted at the outset of the research. The therapy

is most effective for people who present with relatively intact semantic and phonological processing and whose main deficit lies in the mapping between the two levels. Those with impaired processing at these levels are less likely to benefit from this form of therapy.

Interactive therapy, which targets a set of words but encourages use of these in interaction is a further effective means of improving word-finding in aphasia. Again, and not as predicted, the effect of therapy is item-specific, with items not undergoing therapy showing no change or a deterioration in naming. The therapy was effective for all three people with aphasia who took part here.

A main aim of the research was to identify conversation variables which showed adequate reliability and stability to allow their use as outcome measures for therapy. The analysis of reliability and stability led to a set of variables being identified. For two of the participants with aphasia baseline stability was not established so this assessment could not be used to assess outcomes after conversation. For one person (PH) baseline stability was found in a number of variables, and positive changes were found after therapy Phase One. These were not predicted to occur. It was predicted that therapy Phase Two would lead to positive changes however there was no evidence of such changes.

Additional findings from this research relate to specific aspects of the participants' language processing. Assessment of the three participants' language processing enabled their profile to be compared to models of spoken word production. The results of this comparison revealed that no one model of the three considered was able to account for the data from all three people. It may be that further information concerning the language profile of the three individuals is required, in particular there exists no recognized method of testing processing at an intermediate stage between semantics and phonology. And it may also be the case that the models as they stand remain underspecified and thus unable to provide a detailed account of any one individual.

Finally a possible link between the type of cue which was effective for an individual, and the retained function in the corresponding sub-lexical routine, was identified. For KR phonological cues were effective in Phase Two therapy, and auditory-phonological sub-lexical processing was spared, while orthographic cues were not effective in the

context of impaired grapheme-phoneme correspondences. This level of detail is essential in order for a theory of therapy to be articulated.

11.5 OVERALL SUMMARY

In this discussion the strengths and weaknesses of the conversation measure have been outlined and possible alternative ways of measuring lexical retrieval in conversation have been discussed. The results of tests of language processing have been used to identify levels of breakdown in the three individuals and to describe these patterns in terms of models of spoken word production. The effects of the two forms of therapy on the three individuals have been described. Future lines of enquiry have been outlined for both therapy analysis and for measuring conversation.

In this thesis the response of three participants to a lexical therapy has been outlined and some of the possible factors predicting therapy outcome highlighted. A second therapy was also trialled and individual responses again investigated. Encouragingly both forms of therapy were effective to some extent. In the case of Phase One therapy the results provide further evidence for the effectiveness of phonological therapy, which had been deemed many years ago to be less effective than semantic (Howard et al, 1985b). For participants for whom retrieval of the phonological form is the main problem, i.e. those with fairly intact semantic and phonological processing, selection of this therapy would appear to be rational. What is less clear is the relationship between therapy and facilitation. KR responded to the therapy to a greater extent than one would predict from her response to facilitation. It is therefore not possible to predict from the results of facilitation alone whether therapy will work.

The second therapy is more eclectic and the specific effects harder to demarcate. Participants were encouraged to retrieve the words prior to the interaction section, thus some of the elements of Phase One therapy were present. In addition they were asked to use the words in situations approximating to everyday interactions. Again the therapy was effective in terms of gains in picture naming. This is encouraging as items were not exposed to systematic training, and yet still gains were found.

Of importance in future studies is that assessments are thorough, relevant and informative, and that the selected therapy relies on an analysis of multiple factors, not just the hypothesised level of breakdown. Single case descriptions such as those provided here can serve as guidelines for the selection of appropriate therapies for other individuals with aphasia. As more case descriptions are amassed selection criteria for therapy can develop. At present such criteria are fairly rough and are likely to remain so as research in this area proceeds slowly.

In order to identify changes in everyday speech a range of measures is advised, ranging from topic controlled interactions to free conversation. The conversation measure described here can be used to gauge outcomes after therapy. The suggested modifications should improve aspects of reliability and pre-therapy stability. For the person with aphasia participating in therapy, evidence of changes in conversation may provide the most satisfying validation of all their efforts.

REFERENCES

- Alario, F.-X., Schiller, N.O., Domoto-Reilly, K. & Caramazza, A. (2003) The role of phonological and orthographic information in lexical selection. *Brain and Language*, 84, 372-398.
- Albert, M.L. & Sandson, J. (1986) Perseveration in aphasia. *Cortex*, 22, 103-115.
- Albert, M.S., Heller, H.S., & Milberg, W. (1988) Changes in naming ability with age. *Psychology and Aging*, 3, 173-178.
- Albyn Davis, G. (1993) *A survey of adult aphasia and related language disorders* (Second Edition). New Jersey: Prentice-Hall.
- Aten, J., Caliguiri, M., & Holland, A. (1982) The efficacy of functional communication therapy for chronic aphasic patients. *Journal of speech and Hearing Disorders*, 47, 93-96.
- Avila, C., Lambon Ralph, M.A., Parcet, M.-A., Geffner, D., & Gonzalez-Darder, J.-M. (2001) Implicit word cues facilitate impaired naming performance: Evidence from a case of anomia. *Brain and Language*, 79, 185-200.
- Baars, B.J., Motley, M.T. and MacKay, D. (1975). Output editing for lexical status from artificially elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behaviour*, 14, 382-391.
- Badecker, W., Miozzo, M., & Zanuttini, R. (1995) The two stage model of lexical retrieval: evidence from a case of anomia with selective preservation of grammatical gender. *Cognition*, 57, 193-216.
- Barry, C. and McHattie, J. (1991) Depth of semantic processing in picture naming facilitation in aphasic patients. *Paper presented at the British Aphasiology Society Conference. Sheffield, September 1991.*
- Basso, A., Marangolo, P., Piras, F. and Galluzzi, C. (2001) Acquisition of new "words" in normal participants: a suggestion for the treatment of anomia. *Brain and Language*, 77, 45-59.
- Benson, D.F. (1979) *Aphasia, Alexia, and Agraphia*. New York: Churchill Livingstone.
- Berndt, R.S. (1998) Syntactic Aspects of Aphasia. In M.T. Sarno (Ed.) *Acquired Aphasia (Third Edition)*. London: Academic Press
- Berndt, R.S., Burton, M.W., Haendiges, A.N., & Mitchum, C.C. (2002) Production of nouns and verbs in aphasia: Effects of elicitation context. *Aphasiology*, 16, 83-106.
- Best, W. & Nickels, L. (2000) From theory to therapy in aphasia. Where are we now and where to next? *Neuropsychological Rehabilitation*, 10, 231-247.
- Best, W. (1994) Unpublished PhD Thesis, University of London.
- Best, W. (1995) A reverse length effect in dysphasic naming: When elephant is easier than ant. *Cortex*, 31, 637-652.
- Best, W., Herbert, R., Hickin, J., Osborne, F. and Howard, D. (2002) Phonological and orthographic facilitation of word-retrieval in aphasia: immediate & delayed effects. *Aphasiology*, 16, 151-168.
- Best, W.M. (1996). When racquets are baskets but baskets are biscuits, where do the words come from? A single-case study of formal paraphasic errors in aphasia. *Cognitive Neuropsychology*, 13, 443-480.
- Best, W., Howard, D., Bruce, C. and Gatehouse, C. (1997) Cueing the words: A single case study of treatments for anomia. *Neuropsychological Rehabilitation*, 7, 105-141.
- Bird, H., & Franklin, S. (1996) Cinderella revisited: A comparison of fluent and non-fluent aphasic speech. *Journal of Neurolinguistics*, 9, 187-206.

- Bird, H., Howard, D., & Franklin, S. (2001) Why is a verb like an inanimate object? Grammatical category and semantic category deficits. *Brain and Language*, 72, 246-309.
- Bird, H., Lambon Ralph, M.A., Patterson, K., & Hodges, J.R. (2000) The fall and rise of frequency and imageability: How the progression of semantic dementia impacts on noun and verb production in the Cookie Theft description. *Brain and Language*, 73, 17-49.
- Bland, J.M., & Altman, G.G. (1999) Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*, 8, 135-160.
- Blanken, G. (1991) The functional basis of speech automatisms (recurring utterances). *Aphasiology*, 5, 103-127.
- Blanken, G. (1990) Formal paraphasias: a single case study. *Brain and Language*, 38, 534-554.
- Blanken, G. (1998). Lexicalisation in speech production: evidence from form-related word substitutions in aphasia. *Cognitive Neuropsychology*, 15, 321-360.
- Blomert, L., Koster, C., van Mier, H., & Kean, M.-L. (1987) Verbal communication abilities of aphasic patients: The everyday language test. *Aphasiology*, 1, 463-474.
- Boles, L. (1998) Conversational discourse analysis as a method for evaluating progress in aphasia: A case report. *Journal of Communication Disorders*, 31, 261-274.
- Boles, L., & Bombard, T. (1998) Conversational discourse analysis: Appropriate and useful sample sizes. *Aphasiology*, 12, 547-560.
- Booth, S., & Perkins, L. (1999) The use of conversation analysis to guide individualised advice to carers and evaluate change in aphasia: A case study. *Aphasiology*, 13, 283-303.
- Breedin, S., Saffran, E., & Coslett, B. (1994) Reversal of the concreteness effect in a patient with semantic dementia. *Cognitive Neuropsychology*, 11, 617-660.
- Brookshire, R.H., & Nicholas, L.E. (1994) Speech sample size and test-re-test stability of connected speech measures for adults with aphasia. *Journal of Speech and Hearing Research*, 37, 399-407.
- Brown, A.S. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin*, 109, 204-223.
- Brown, R. & McNeill, D. (1966) The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behaviour*, 5, 325-337.
- Bruce, C. and Howard, D. (1987) Computer generated cues: an effective aid for naming in aphasia. *British Journal of Disorders of Communication*, 22, 191-201.
- Bruce, C. and Howard, D. (1988). Why don't Broca's aphasics cue themselves? An investigation of phonemic cueing and tip-of-the-tongue information. *Neuropsychologia*, 26, 253-264.
- Burke, D.M., MacKay, D.G., Worthley, J.A., & Wade, E. (1991) On the tip of the tongue: What causes word finding failures in young and older adults? *Journal of Memory and Language*, 30, 542-579.
- Butterworth, B. (1980) Evidence from pauses. In B.L. Butterworth (Ed.) *Language Production, Volume I: Speech and Talk*. London: Academic Press.
- Butterworth, B. (1989). Lexical access in speech production. In W. Marslen-Wilson (ed.), *Lexical Representation and Process*. Cambridge, MA: MIT.
- Butterworth, B. (1992). Disorders of phonological encoding. *Cognition*, 42, 261-286.
- Butterworth, B., Howard, D., & McLoughlin, P. (1984) The semantic deficit in aphasia: The relationship between semantic errors in auditory comprehension and picture naming. *Neuropsychologia*, 22, 409-426.

- Caramazza, A. & Miozzo, M. (1997) The relation between syntactic and phonological knowledge in lexical access: evidence from the 'tip-of-the-tongue' phenomenon. *Cognition*, 64, 309-343.
- Caramazza, A. & Hillis, A. E. (1993) For a theory of rehabilitation. *Neuropsychological Rehabilitation*, 3, 217-234.
- Caramazza, A. & Hillis, A.E. (1990) Where do semantic errors come from? *Cortex*, 26, 95-122.
- Caramazza, A. (1989) Cognitive neuropsychology and rehabilitation: an unfulfilled promise. In Seron, X. and Deloche, G. (1989) (Eds.) *Cognitive Approaches in Neuropsychological Rehabilitation*. London: LEA.
- Caramazza, A. (1997) How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, 14, 1, 177-208.
- Caramazza, A. and Hillis, A.E. (1991). Lexical organisation of nouns and verbs in the brain. *Nature*, 349, 788-790.
- Carletta, J. (1996) Assessing agreement on classification tasks: The Kappa statistic. *Computational Linguistics*, 22, 1-6.
- Cave, C.B. (1997). Very long-lasting priming in picture naming. *Psychological Science*, 8, 322-325.
- CELEX: Centre for Lexical Information (1993) *The Celex Database*. Nijmegen, The Netherlands: Max Plank Institute for Psycholinguistics.
- Clark, H., and Schaefer, E. (1987) Collaborating on contributions to conversation. *Language and Cognitive Processes*, 2, 19-41.
- Coelho, C.A., McHugh, R.E., & Boyle, M. (2000) Semantic feature analysis as a treatment for aphasic dysnomia: A replication. *Aphasiology*, 14, 133-142.
- Cohen, G., & Faulkner, D. (1986) Memory for proper names: Age differences in retrieval. *British Journal of Developmental Psychology*, 4, 187-197.
- Cohen, J. (1960) A coefficient of agreement for nominal scales. *Psychological Measures*, 20, 37-46.
- Cohen, L., Bolger, F., Timsit, S., & Chermann, J.F. (1994) Anomia for proper-names after left thalamic infarct. *Journal of Neurology, Neurosurgery, and Psychiatry*, 57, 1283-1284.
- Coltheart, M. (1980) The semantic error: types and theories. In M. Coltheart, K. Patterson, & J. Marshall (Eds.) *Deep Dyslexia*. London: Routledge and Kegan Paul.
- Coltheart, M. (1981) The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology*, 31, 489-507.
- Coltheart, M. (2000) Deep dyslexia is right-hemisphere reading. *Brain and Language*, 71, 299-309
- Coltheart, M., Inglis, L., Cupples, L., Michie, P., Bates, A., Budd, B. (1998) A semantic subsystem of visual attributes. *Neurocase*, 4, 353-370.
- Coltheart, M., Patterson, K.E., Marshall, J.C. (Eds.) (1980) *Deep Dyslexia*. London: Routledge & Kegan Paul.
- Comrie, P., Mackenzie, C., & McCall, J. (2001) The influence of acquired dysarthria on conversational turn-taking. *Clinical Linguistics and Phonetics*, 15, 383-398.
- Craik, F.I.M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 104, 268-294.
- Crockford, C. and Lesser, R. (1994) Assessing functional communication in aphasia: Clinical utility and time demands of three methods. *European Journal of Disorders of Communication*, 29, 165-182.
- Crutch, S.J., & Warrington, E.K. (2003) Preservation of propositional speech in a pure anomia: The importance of an abstract vocabulary. *Neurocase*, 9, 465-481.
- Cutler, A. (1981). The reliability of speech error data. *Linguistics*, 19, 561-582.

- Davis, A. and Pring, T. (1991) Therapy for word-finding deficits: more on the effects of semantic and phonological approaches to treatment with dysphasic patients. *Neuropsychological Rehabilitation*, 1 135-145.
- Davis, G.A., & Wilcox, M.J. (1985). *Adult Aphasia Rehabilitation: Applied Pragmatics*. College Hill Press: San Diego, CA.
- de Partz, M.P. (1986) Re-education of a deep dyslexic patient: rationale of the method and results. *Cognitive Neuropsychology*, 3, 149-177.
- Dell, G.S. & O'Sheaghda, P.G. (1991) Mediated and convergent lexical priming in language production. A comment on Levelt et al (1991). *Psychological Review*, 98, 604-614.
- Dell, G.S. & Reich, P.A. (1981) Stages in sentence production: An analysis of speech error data. *Journal of Verbal Learning and Verbal Behaviour*, 20, 611-629.
- Dell, G.S. (1986) A spreading activation theory of retrieval in language production. *Psychological Review*, 93, 283-321.
- Dell, G.S. (1988) The retrieval of phonological forms in production: Tests of predictions of a connectionist model. *Journal of Memory and Language*, 27, 124-142.
- Dell, G.S. (1989) The retrieval of phonological forms in production: Tests of predictions from a connectionist model. In W. Marslen-Wilson (ed.), *Lexical Representation and Process*. Cambridge, MA: MIT.
- Dell, G.S. (1995) Speaking and Misspeaking. In L.R. Gleitman & M. Liberman (Eds.) *An Invitation to Cognitive Science Volume One: Language*. Cambridge MA: MIT Press.
- Dell, G.S., Juliano, C., & Govindjee, A. (1993) Structure and content in language production: A theory of frame constraints in phonological speech errors. *Cognitive Science*, 17, 149-196.
- Dell, G.S., Lawler, E.N., Harris, H.D., & Gordon, J.K. (2004) Models of errors of omission in aphasic naming. *Cognitive Neuropsychology*, 21, 125-145.
- Dell, G.S., Schwartz, M.F., Martin, N., Saffran, E.M. & Gagnon, D.A. (1997) Lexical access in normal and aphasic speech. *Psychological Review*, 104, 801-838.
- Deloche, G., Hannequin, D., Dordain, M., Perrier, D., Pichard, B., Quint, S., Metz-Lutz, M.-N., Kremin, H., & Cardebat, D. (1996) Picture confrontation oral naming: Performance differences between aphasics and normals. *Brain and Language*, 53, 105-120.
- Doyle, P.J., McNeil, M.R., Park, G., Goda, A., Rubenstein, E., Spencer, K., Carroll, B., Lustig, A., & Szwarc, L. (2000) Linguistic validation of four parallel forms of a story retelling procedure. *Aphasiology*, 14, 537-549.
- Ellis, A.W. and Young, A.W. (1988). *Human Cognitive Neuropsychology*. London: Lawrence Erlbaum Associates.
- Ellis, A.W., Franklin, S. & Crerar, A. (1994) Cognitive neuropsychology and the remediation of disorders of spoken language. In M.J.Riddoch and G.W. Humphreys (Eds.) *Cognitive neuropsychology and cognitive rehabilitation*. Hove UK: Lawrence Erlbaum.
- Ellis, A.W., Lum, C., & Lambon-Ralph, M.A. (1996) The use of regression techniques for the analysis of single case aphasic data. *Journal of Neurolinguistics*, 9, 1-10.
- Eysenck, M. & Keane, M.T. (2000) *Cognitive Psychology A Student's Handbook*. London: Taylor and Francis.
- Fay, D. and Cutler, A. (1977). Malapropisms and the structure of the mental lexicon. *Linguistic Inquiry*, 8, 505-520.

- Ferguson, A. (1994) The influence of aphasia, familiarity and activity on conversation repair. *Aphasiology*, 8, 143-157.
- Ferguson, A. (1996) Describing competence in aphasic normal conversation. *Clinical Linguistics and Phonetics*, 10, 55-63.
- Ferguson, A.J. (1992) Interpersonal aspects of aphasic conversation. *Journal of Neurolinguistics*, 7, 277-294.
- Feyereisen, P., Van Der Borgh, F., & Seron, X. (1988) The operativity effect in naming. A re-analysis. *Neuropsychologia*, 26, 401-415.
- Fink, R., Brecher, A., Schwartz, M., & Robey, R. (2002) A computer implemented protocol for the treatment of naming disorders: Evaluation of clinician-guided and partially self-guided instruction. *Aphasiology*, 16, 1061-1086.
- Flanagan, O. (1996) *The Science of the Mind* (Second Edition). Cambridge MA: MIT Press.
- Foygel, D., & Dell, G.S. (2000) Models of impaired lexical access in speech production. *Journal of Memory and Language*, 43, 182-216.
- Franklin, S.E. (1997) Designing single case treatment studies for aphasic patients. *Neuropsychological Rehabilitation*, 7, 401-418.
- Franklin, S., Turner, J., and Ellis, A. W. (1992). *ADA (Action for Dysphasic Adults) Comprehension Battery*. London: Action for Dysphasic Adults.
- Franklin, S.E., Buerk, F., & Howard, D. (2002) Generalised improvement in speech production for a subject with reproduction conduction aphasia. *Aphasiology*, 16, 1087-1114.
- Freud, S. (1958) *Psychopathology of Everyday Life*. Translated by A.A. Brill. New York: New American Library.
- Gainotti, G. (1976) The relationship between semantic impairment in comprehension and naming in aphasic patients. *British Journal of Disorders of Communication*, 11, 57-61.
- Gainotti, G., Miceli, G., Caltagirone, C., Silveri, M.C., & Masullo, C. (1981) The relationship between type of naming error and semantic-lexical discrimination in aphasia. *Cortex*, 17, 401-410.
- Gainotti, G., Silveri, M.C., Villa, G. and Miceli, G. (1986) Anomia with and without lexical comprehension disorders. *Brain and Language*, 29, 18-33.
- Gardner, H. (1973) The contribution of operativity to naming capacity in aphasic patients. *Neuropsychologia*, 11, 213-220.
- Gardner, W. (1995) On the reliability of sequential data: Measurement, meaning and correction. In J.M. Gottman (Ed.) *The analysis of change*. Mahwah, NJ: Erlbaum.
- Garnham, A.R.C., Shillcock, G.D.A., Brown, A.I.D., & Cutler, A. (1981) Slips of the tongue in the London-Lund corpus of spontaneous conversation. *Linguistics*, 19, 805-817.
- Garrett, M.F. (1975). The analysis of sentence production. In G.H. Bower (Ed.) *The psychology of learning and motivation*. New York: Academic Press.
- Garrett, M.F. (1980). Levels of processing in sentence production. In B.L. Butterworth (Ed.), *Language Production, Volume 1: Speech and Talk*. London: Academic Press.
- Gernsbacher, M.A. (1984) resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General*, 113, 256-281.
- Gold, B.T., & Kertesz, A. (2001) Phonologically related repetition disorder: A case study. *Brain and Language*, 77, 241-265.
- Goodglass, H. (1993) *Understanding Aphasia*. San Diego: Academic Press.
- Goodglass, H., & Wingfield, A. (1997) *Anomia*. San Diego California: Academic Press.

- Goodglass, H., Kaplan, E., Weintraub, S., & Ackerman, N. (1976) The tip of the tongue phenomenon in aphasia. *Cortex*, 12, 145-153.
- Goodwin, C. (1981) *Conversational Organisation*. New York: Academic Press.
- Goulet, P., Ska, B., & Kahn, H.J. (1994) Is there a decline in picture naming with advancing age? *Journal of Speech and Hearing Research*, 37, 629-644.
- Graham, K.S., Patterson, K., & Hodges, J.R. (1995) Progressive pure anomia: Insufficient activation of phonology by meaning. *Neurocase*, 1, 25-38.
- Greenhouse, J.B., Bromberg, J.A., & Fromm, D. (1995) An introduction to logistic regression with an application to the analysis of language recovery after a stroke. *Journal of Communication Disorders*, 28, 229-246.
- Greenwald, M.L., Raymer, A.M., Richardson, M.E., & Rothi, L.G.J. (1995) Contrasting treatments for severe impairments of picture naming. *Neuropsychological Rehabilitation*, 5, 17-49.
- Harley, T. A. (1996). Connectionist modelling of the recovery of language functions following brain damage. *Brain and Language*, 52, 7-24.
- Harley, T.A. (1996) Connectionist modeling of the recovery of language functions after brain damage. *Brain and Language*, 52, 7-24.
- Harley, T.A. (1995) Connectionist model of anomia: A comment on Nickels. *Language and Cognitive Processes*, 10, 47-58.
- Hebb, D.O. (1949) *The Organization of Behavior: A Neuropsychological Theory*. New York: Wiley.
- Helasvuo, M.-L. (2001) Emerging syntax for interaction: Noun phrases and clauses as a syntactic resource for interaction. In M. Selting & E. Couper-Kuhlen (eds.) *Studies in Interactional Linguistics*. Amsterdam: John Benjamins.
- Helm-Estabrooks, N. (1997) Treatment of Aphasic Naming Problems. In H. Goodglass and A. Wingfield (Eds.) *Anomia*. San Diego: Academic Press.
- Henaff Gonon, M.A., Bruckert, R. and Michel, F. (1989). Lexicalization in an anomic patient. *Neuropsychologia*, 27, 391-407.
- Herbert, R. (1998) A cognitive neuropsychological approach to bilingual aphasia. Unpublished MSc thesis, University College London.
- Herbert, R., Best, W., Hickin, J., Howard, D. & Osborne, F. (2003) Combining lexical and interactional approaches to therapy for word finding deficits in aphasia. *Aphasiology*, 17, 1163-1186.
- Heritage, J. (1999) Conversation analysis at century's end: Practices of talk-in-interaction, their distributions, and their outcomes. *Research on Language and Social Interaction*, 32, 69-76.
- Hickin, J., Best, W., Herbert, R., Howard, D. & Osborne, F. (2002) Phonological therapy for word finding difficulties: a re-evaluation. *Aphasiology*, 16, 981-999.
- Hillis, A. (1989) Efficacy and generalization of treatment for aphasic naming errors. *Archives of Physical Medicine and Rehabilitation*, 70, 632-636.
- Hillis, A.E. & Caramazza, A. (1995) The compositionality of lexical semantic representations: Clues from semantic errors in object naming. *Memory*, 3, 333-358.
- Hillis, A.E. and Caramazza, A. (1994) Theories of lexical processing and rehabilitation of lexical deficits. In Riddoch, M.J. and Humphreys, G.W. (1994) (Eds.) *Cognitive Neuropsychology and Cognitive Rehabilitation*. Hove: LEA.
- Hillis, A.E., Boatman, D., Hart, J., & Gordon, B. (1999) Making sense out of jargon: A neurolinguistic and computational account of jargon aphasia. *Neurology*, 53, 1813-1824
- Hillis, A.E., Rapp, B.C., Romani, C. and Caramazza, A. (1990). Selective impairment of semantics in lexical processing. *Cognitive Neuropsychology*, 7, 191-243.

- Hirsh, K.W., & Ellis, A.W. (1994) Age of acquisition and lexical processing in aphasia: A case study. *Cognitive Neuropsychology*, 11, 435-458.
- Hirsh, K.W., & Funnell, E. (1995) Those old familiar things: Age of acquisition, familiarity, and lexical access in progressive aphasia. *Journal of Neurolinguistics*, 9, 23-32.
- Holland, A. L. (1991) Pragmatic aspects of intervention in aphasia. *Journal of Neurolinguistics*, 6, 197-211.
- Holland, A.L. (1980) *Communicative abilities in daily living*. Baltimore: University Park Press.
- Holland, A.L. (1994) Cognitive neuropsychological theory and treatment for aphasia: Exploring the strengths and limitations. *Clinical Aphasiology*, 22, 275-282.
- Howard, D. & Harding, D. (1998). Self-cueing of word retrieval by a woman with aphasia: why a letter board works. *Aphasiology*, 12, 399-420.
- Howard, D. & Orchard-Lisle, V. (1984) On the origin of semantic errors in naming: Evidence from the case of a global aphasic. *Cognitive Neuropsychology*, 1, 163-90.
- Howard, D. (1986) Beyond randomised controlled trials: the case for effective case studies of the effects of treatment in aphasia. *British Journal of Disorders of Communication*, 21, 89-102.
- Howard, D. (1995) Lexical anomia (or the case of the missing entries). *Quarterly Journal of Experimental Psychology*, 48A, 999-1023.
- Howard, D. (2000) Cognitive neuropsychology and aphasia therapy: the case of word retrieval. In I. Papathanasiou (Ed) *Acquired Neurogenic Communication Disorders: A Clinical Perspective*. London: Whurr.
- Howard, D. and Franklin, S. (1988). *Missing the meaning?* Cambridge, Massachusetts: M.I.T. Press.
- Howard, D., Best, W., Bruce, C., & Gatehouse, C. (1995) Operativity and animacy effects in aphasic naming. *European Journal of Disorders of Communication*, 30, 286-302.
- Howard, D., & Patterson, K.E. (1992) *The Pyramids and Palm Trees Test*. Bury St Edmunds: Thames Valley Test Company.
- Howard, D., Patterson, K.E., Franklin, S.E., Morton, J., & Orchard-Lisle, V.M. (1984) Variability and consistency in picture naming by aphasic patients. *Advances in Neurology*, 42, 263-276.
- Howard, D., Patterson, K.E., Franklin, S.E., Orchard-Lisle, V., & Morton, J. (1985b) The treatment of word retrieval deficits in aphasia: a comparison of two therapy methods. *Brain*, 108, 817-829.
- Howard, D., Patterson, K.E., Franklin, S.E., Orchard-Lisle, V.M., & Morton, J. (1985a) The facilitation of picture naming in aphasia. *Cognitive Neuropsychology*, 2, 49-80.
- Hutchby, I., & Wooffitt, R. (1998) *Conversation Analysis*, Cambridge, UK: Polity Press.
- Jefferson, G. (1984) On the organization of laughter in talk about troubles. In J.M. Atkinson and J. Heritage (Eds.) *Structures of Social Action: Studies in Conversation Analysis*. Cambridge: Cambridge University Press.
- Jescheniak, J.D. & Levelt, W.J.M. (1994) Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning Memory and Cognition*, 20, 824-843.
- Jescheniak, J.D. (1994) Word frequency effects in speech production. Unpublished doctoral dissertation. Nijmegen University.

- Jokel, R. & Rochon, E. (1996) Treatment for an aphasic naming impairment: When phonology met orthography. *Brain and Cognition*, 32, 299-301.
- Kagan, A. (1998) Supported conversation for adults with aphasia: Methods and resources for training conversation partners. *Aphasiology*, 12, 811-864.
- Katz, R., & Goodglass, H. (1990) Deep dysphasia: analysis of a rare form of repetition disorder. *Brain and Language*, 39, 153-185.
- Kay, J. & Ellis, A.W. (1987) A cognitive neuropsychological case study of anomia: Implications for psychological models of word retrieval. *Brain*, 110, 613-29.
- Kay, J., Lesser, R. & Coltheart, M. (1992). *Psycholinguistic Assessment of Language Processing in Aphasia*. London: Lawrence Erlbaum.
- Keefe, K.A. (1995) Applying basic neuroscience to aphasia therapy: What the animals are telling us. *American Journal of Speech-Language Pathology*, 4, 88-93.
- Kempen, G. and Huijbers, P. (1983). The lexicalisation process in sentence production and naming: Indirect election of words. *Cognition*, 14, 185-209.
- Kennedy, M.R.T., Strand, E.A., Burton, W., & Peterson, C. (1994) Analysis of first encounter conversation of right-hemisphere-damaged adults. In M.L. Lemme (Ed.) *Clinical Aphasiology, Volume 22*. Austin: Pro-Ed.
- Kohn, S.E., & Goodglass, H. (1985) Picture naming in aphasia. *Brain and Language*, 24, 266-283.
- Kolb, B. (1996) *Brain Plasticity and Behaviour*. Hillsdale, NJ: Erlbaum.
- Kulke, F., & Blanken, G. (2001) Phonological and syntactic influences on semantic misnamings in aphasia. *Aphasiology*, 15, 3-15.
- Laakso, M., & Klippi, A. (1999) A closer look at the 'hint and guess' sequences in aphasic conversation. *Aphasiology*, 13, 345-363.
- Lambon Ralph, M. A., Sage, K. & Roberts, J. (2000) Classical anomia: a neuropsychological perspective on speech production. *Neuropsychologia*, 38, 186-202.
- Lambon Ralph, M., Cippolotti, L. and Patterson, K. (1999) Oral naming and oral reading: do they speak the same language? *Cognitive Neuropsychology*, 16, 157-169.
- Lambon Ralph, M.A. (1998) Distributed versus localist representations: Evidence from a study of item consistency in a case of classical anomia. *Brain and Language*, 64, 339-360.
- Landis, J.R. & Koch, G.G. (1977) The measurement of observer agreement for categorical data. *Biometrics*, 33, 159-174.
- Le Dorze, G. & Nespoulous, J.L. (1989) Anomia in moderate aphasia: Problems in accessing the lexical representation. *Brain and Language*, 37, 381-400.
- Le Dorze, G., Boulay, N., Gaudreau, J., Brassard, C. (1994) The contrasting effects of a semantic versus a formal-semantic technique for the facilitation of naming in a case of anomia. *Aphasiology*, 8, 127-141.
- Lesser, R. (1989) Some issues in the neuropsychological rehabilitation of anomia. In X. Seron and G. Deloche (Eds.) *Cognitive Approaches in Neuropsychological Rehabilitation*. Hillsdale NJ: Lawrence Erlbaum Associates.
- Lesser, R. and Algar, L. (1995) Towards combining the cognitive neuropsychological and the pragmatic in aphasia therapy. *Neuropsychological Rehabilitation*, 5, 67-92.
- Lesser, R. and Milroy, L. (1993) *Linguistics and Aphasia: Psycholinguistic and pragmatic aspects of intervention*. London: Longman.
- Levelt, W.J.M. (1983) Monitoring and self-repair in speech. *Cognition*, 14, 41-104.
- Levelt, W.J.M. (1989) *Speaking: From intention to articulation*. Cambridge MA: MIT Press.

- Levelt, W.J.M. (1992) Accessing words in speech production: Stages, processes and representations. *Cognition*, 42, 122.
- Levelt, W.J.M., Roelofs, A., & Meyer, A.S. (1999) A theory of lexical access in speech production. *Behavioural and Brain Sciences*, 22, 1-75.
- Levelt, W.J.M., Schriefers, H., Vorberg, D., Meyer, A.S., Pechmann, T. and Havinga, J. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review*, 98, 122-142.
- Levinson, S. C. (1983) *Pragmatics*. Cambridge UK: Cambridge University Press.
- Lindsay, J., & Wilkinson, R. (1999) Repair sequences in aphasic talk: A comparison of aphasic-speech and language therapist and aphasic-spouse conversations. *Aphasiology*, 13, 305-326.
- Lomas, J., Pickard, L., Bester, S., Elbard, H., Finlayson, A., & Zoghaib, C. (1989) The Communicative Effectiveness Index. Development and psychometric evaluation of a functional communication measure for adult aphasia. *Journal of Speech and Hearing Disorders*, 54, 113-124.
- Lubinski, R., Duchan, J., & Weitzner-Lin, B. (1980) Analysis of breakdowns and repairs in aphasic adult communication. In R.H. Brookshire (Ed.) *Clinical Aphasiology Conference Proceedings* (pp. 111-116). Minneapolis: BRK.
- MacKay, D.G. (1972) The structure of words and syllables: Evidence from errors in speech. *Cognitive Psychology*, 3, 210-227.
- Marr, D. (1982) *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: W.H. Freeman.
- Marshall, J., Pound, C., White-Thomson, M. and Pring, T. (1990) The use of picture/word matching tasks to assist word retrieval in aphasic patients. *Aphasiology*, 4, 167-84.
- Marshall, J., Pring, T., Chiat, S., & Robson, J. (1996) Calling a salad a federation: An investigation of semantic jargon.1. Nouns. *Journal of Neurolinguistics*, 9, 237-250.
- Marshall, J., Pring, T., Chiat, S., & Robson, J. (2001) When ottoman is esier than chair: An inverse frequency effect in jargon aphasia. *Cortex*, 37, 33-53.
- Marshall, J., Robson, J., Pring, T., & Chiat, S. (1998) Why does monitoring fail in jargon aphasia? Comprehension, judgement and therapy evidence. *Brain and Language*, 63, 79-107.
- Marshall, J.C. & Newcombe, F. (1973). Patterns of paralexia: a psycholinguistic approach. *Journal of Psycholinguistic Research*, 2:175-199.
- Martin, N. & Saffran, E.M. (1992). A computational account of deep dysphasia: Evidence from a single case study. *Brain and Language*, 43, 240-274.
- Martin, N., Dell, G.S., Saffran, E.M. & Schwartz, M.F. (1994). Origins of paraphasias in deep dysphasia: Testing the consequences of a decay impairment to an interactive spreading activation model of lexical retrieval. *Brain and Language*, 47, 609-660.
- Martin, N., Gagnon, D.A., Schwartz, M.F., Dell, G.S. & Saffran., E.M. (1996). Phonological facilitation of semantic errors in normal and aphasic speakers. *Language and Cognitive Processes*, 11, 257-282.
- Martin, N., Loach, R., Brecher, A., & Lowery, J. (1998) Lexical retrieval mechanisms underlying whole word perseveration errors in anomic aphasia. *Aphasiology*, 12, 319-333.
- Martin, N., Saffran, E.M., Dell, G.S., & Schwartz, M. (1994) Origins of paraphasias in deep dysphasia: Testing the consequences of a decay impairment to an interactive spreading activation model of lexical retrieval. *Brain and Language*, 47, 609-660.

- Martin, N., Weisberg, R.W., & Saffran, E.M. (1989) Variables influencing the occurrence of naming errors: implications for models of lexical retrieval. *Journal of Memory and Language*, 28, 462-485.
- Meyer, A. S. & Schriefers, H. (1991) Phonological facilitation in picture-word interference experiments: effects of stimulus onset asynchrony and types of interfering stimuli. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 17: 1146-60.
- Meyer, A.S. (1990). The time course of phonological encoding in language production: the encoding of successive syllables of a word. *Journal of Memory and Language*, 29, 524-545.
- Meyer, A.S., & Bock, K. (1992) The tip of the tongue phenomenon: Blocking or partial activation? *Memory and Cognition*, 20, 715-726.
- Miceli, G., Amitrano, A., Capasso, R., & Caramazza, A. (1996) The remediation of anomia resulting from output lexical damage: analysis of two cases. *Brain and Language*, 52: 150-174.
- Milroy, L. and Perkins, L. (1992) Repair strategies in aphasic discourse: towards a collaborative model. *Clinical Linguistics and Phonetics*, 6, 27-40.
- Miozzo, M. & Caramazza, A. (1997). Retrieval of lexical-syntactic features in tip-of-the-tongue state. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 23, 1-14.
- Monsell, S. (1987). On the relation between lexical input and output pathways for speech. In A. Allport, D. MacKay, W. Prinz and E. Scheerer (Eds.) *Language perception and production: relationships between listening, speaking, reading and writing*. London: Academic Press.
- Monsell, S., Mathews, G. H., and Miller, D. C. (1992) Repetition of lexicalisation across languages: a further test of the locus of priming. *Quarterly Journal of Experimental Psychology*, 44A, 763-783.
- Morton, J. (1980) The logogen model and orthographic structure. In U. Frith (Ed.) *Cognitive processes in spelling*. London: Academic Press.
- Morton, J. (1985). Naming. In S.K. Newman and R. Epstein (Eds.) *Current perspectives in dysphasia*. Edinburgh: Churchill Livingstone.
- Nettleton, J. and Lesser, R. (1991) Therapy for naming difficulties in aphasia: Application of a cognitive neuropsychological model. *Journal of Neurolinguistics*, 6, 139-157.
- Nicholas, L.E. & Brookshire, R.H. (1993) A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. *Journal of Speech and Hearing Research*, 36, 338-350.
- Nicholas, L.E., Brookshire, R.H., MacLennan, D.L., Schumacher, J.G., & Porrazzo, S.A. (1989) The Boston Naming Test: Revised administration and scoring procedures and normative information for non-brain-damaged adults. In T.E. Prescott (Ed.) *Clinical Aphasiology, Volume 18*. Austin, Texas: Pro-Ed.
- Nicholas, M., Obler, L.K., Albert, M.L. & Goodglass, H. (1985) Lexical retrieval in healthy aging. *Cortex*, 21, 595-606.
- Nickels, L. (1992) *European Naming Test* (unpublished).
- Nickels, L. (2002a) Therapy for naming disorders: Revisiting, revising and reviewing. *Aphasiology*, 16, 935-979.
- Nickels, L. (2002b) Improving word finding: Practice makes (closer to) perfect? *Aphasiology*, 16, 1047-1060.
- Nickels, L.A. (1992) The autocue? Self-generated phonemic cues in the treatment of a disorder of reading and naming. *Cognitive Neuropsychology*, 9, 155-182.

- Nickels, L.A. (1995) Getting it right? Using aphasic naming errors to evaluate theoretical models of spoken word production. *Language and Cognitive Processes*, 10, 13-45.
- Nickels, L. A. (1997). *Spoken word production and its breakdown in aphasia*. Hove, UK: Psychology Press.
- Nickels, L.A., & Best, W. (1996a) Therapy for naming disorders (Part I): principles, puzzles and progress. *Aphasiology*, 10, 21-47.
- Nickels, L.A., & Best, W. (1996b) Therapy for naming disorders (Part II): specifics, surprises and suggestions. *Aphasiology*, 10, 109-136.
- Nickels, L.A., & Howard, D. (1994) A frequent occurrence? factors affecting the production of semantic errors in aphasic naming. *Cognitive Neuropsychology*, 11, 289-320.
- Nickels, L.A., & Howard, D. (1995) Phonological errors in aphasic naming; comprehension, monitoring and lexicality. *Cortex*, 31, 209-237.
- Nickels, L.A. & Howard, D. (2004) Dissociating effects of number of phonemes, number of syllables, and syllabic complexity on word production in aphasia: It's the number of phonemes that counts. *Cognitive Neuropsychology*, 21, 57-78.
- Oelschlaeger, M.L. & Damico, J.S. (1998) Joint productions as a conversational strategy in aphasia. *Clinical Linguistics and Phonetics*, 12, 459-480.
- Oelschlaeger, M.L. (1999) Participation of a conversation partner in the word searches of a person with aphasia. *American Journal of Speech-Language Pathology*, 8, 62-71.
- Oldfield, R.C. and Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, 17, 273-281.
- Parr, S., Byng, S., & Gilpin, S. (1997) *Talking about aphasia*. Buckingham, UK: Open University Press.
- Parr, S., Duchan, J., & Pound, C. (2003) *Aphasia Inside Out*. Buckingham UK: Open University Press.
- Pashek, G.V., & Tompkins, C.A. (2002) Context and word class influences on lexical retrieval in aphasia. *Aphasiology*, 16, 261-286.
- Patterson, K.E., & Shewell, C. (1987) *The Cognitive Neuropsychology of Language*. Hove: Lawrence Erlbaum Associates.
- Patterson, K.E., Purell, C., & Morton, J. (1983) The facilitation of naming in aphasia. In C. Code, & D.J. Mueller (Eds.) *Aphasia therapy*. London: Arnold.
- Pease, D.M., & Goodglass, H. (1978) The effects of cuing on picture naming in aphasia. *Cortex*, 14, 178-189.
- Perkins, L. (1995) Applying conversation analysis to aphasia: Clinical implications and analytic issues. *European Journal of Disorders of Communication*, 30, 372-383.
- Perkins, L., Crisp, J., & Walshaw, D. (1999) Exploring conversation analysis as an assessment tool for aphasia: The issue of reliability. *Aphasiology*, 3, 259-281.
- Peterson, R. R. & Savoy, P. (1998) Lexical selection and phonological encoding during language production: evidence for cascaded processing. *Journal of Experimental Psychology; Learning Memory and Cognition*, 24: 539-57.
- Plaut, D. and Shallice, T. (1993). Deep Dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology*, 10, 377-500.
- Plaut, D.C., McClelland, J.L., Seidenberg, M.S., & Patterson, K. (1996) Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, 103, 56-105.
- Porch, B.E. (1967) *Porch Index of Communicative Ability, Volume I, Theory and development*. Palo Alto, CA: Consulting Psychologists Press.

- Pring, T.R. (1986) Evaluating the effects of speech therapy for aphasics: developing the single case methodology. *British Journal of Disorders of Communication*, 21, 103-115.
- Pring, T.R., White-Thomson, M., Pound, C., Marshall, J., & Davis, A. (1990) Picture/word matching tasks and word retrieval: some follow-up data and second thoughts. *Aphasiology*, 4, 479-483.
- Pulvermuller, F., & Roth, V.M. (1991) Communicative aphasia treatment as a further development of PACE therapy. *Aphasiology*, 5, 39-50.
- Ramsberger, G., & Rende, B. (2002) Measuring transactional success in the conversation of people with aphasia. *Aphasiology*, 16, 337-353.
- Rapp, B. & Caramazza, A. (1997) The modality-specific organization of grammatical categories: Evidence from impaired spoken and written sentence production. *Brain and Language*, 56, 248-286.
- Rapp, B. & Goldrick, M. (2000) Discreteness and interactivity in spoken word production. *Psychological Review*, 107, 460-499.
- Rapp, B., Benzig, L., & Caramazza, A. (1995) Evidence regarding the independence of orthographic and phonological representations. *Brain and Cognition*, 28, 205-205.
- Raven, J.C. (1976) *Coloured Progressive Matrices*. Oxford: Oxford Psychologists Press.
- Raymer, A. M. Thompson, C. K., Jacobs, B. and Le Grand, H.R. (1993) Phonological treatment of naming deficits in aphasia: Model-based generalization analysis. *Aphasiology*, 7, 27-53.
- Riddoch, M.J. & Humphreys, G.W. (1992) *BORB: Birmingham Object Recognition Battery*. Hove UK: Lawrence Erlbaum Associates.
- Riddoch, M.J. and Humphreys, G.W. (1994) (Eds.) *Cognitive Neuropsychology and Cognitive Rehabilitation*. Hove: LEA.
- Rinnert, C., & Whitaker, H.A. (1973) Semantic confusions by aphasic patients. *Cortex*, 9, 56-81.
- Roach, A., Schwartz, M.F., Martin, N., Grewal, R.S., & Brecher, A. (1996) The Philadelphia Naming Test: Scoring and rationale. *Clinical Aphasiology*, 24, 121-133.
- Robertson, I.H. & Murre, J.M.J. (1999) Rehabilitation of brain damage: Brain plasticity and principles of guided recovery. *Psychological Bulletin*, 125, 544-575.
- Robson, J., Pring, T., Marshall, J., Morrison, S., & Chiat, S. (1998) Written communication in undifferentiated jargon aphasia: a therapy study. *International Journal of Language and Communication Disorders*, 33, 305-328.
- Rochford, G. & Williams, M. (1965) Studies in the development and breakdown of the use of names, IV: The effects of word frequency. *Journal of Neurology, Neurosurgery, and Psychiatry*, 28, 407-413.
- Roelofs, A. (1996) Serial order in planning the production of successive morphemes of a word. *Journal of Memory and Language*, 35, 854-876.
- Rumel, W., & Caramazza, A. (2000) An evaluation of a computational model of lexical access: Comment on Dell et al. (1997). *Psychological Review*, 107, 609-634.
- Sacks, H. (1992) *Lectures on conversation: Volumes I and II*. Oxford: Blackwell.
- Sacks, H., Schegloff, E., & Jefferson, G. (1974) A simplest systematics for the organisation of turn-taking in conversation. *Language*, 50, 696-735.
- Sarno, M.T. (1969) The functional communication profile manual of directions. *Rehabilitation Monograph*, 42, New York University Medical Center.
- Schegloff, E. (1982) Discourse as an interactional achievement: some uses of 'uh huh' and other things that come between sentences. In D. Tannen (Ed.) *Georgetown*

University Roundtable on Languages and Linguistics 93. Washington DC: Georgetown University Press.

- Schegloff, E. (1988) Presequences and indirection: applying speech act theory to ordinary conversation. *Journal of Pragmatics*, 12, 55-62.
- Schegloff, E.A., Jefferson, G., & Sacks, H. (1977) The preference for self-correction in the organisation of repair in conversation. *Language*, 53, 361-382.
- Schlenk, K.-J., Huber, W., and Willmes, K. (1987) "Prepairs" and Repairs: different monitoring functions in aphasic language production. *Brain and Language*, 30, 226-244.
- Schriefers, H. (1993) Syntactic processes in the production of noun phrases. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 19, 841-850.
- Schriefers, H., Meyer, A.S. & Levelt, W.J.M. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language*, 29, 86-102.
- Schwartz, M.F., & Brecher, A. (2000) A model-driven analysis of severity, response characteristics, and partial recovery in aphasics' picture naming. *Brain and Language*, 73, 62-91.
- Schwartz, M.F., Saffran, E.M., Bloch, D.E., & Dell, G.S. (1994) Disordered speech production in aphasic and normal speakers. *Brain and Language*, 47, 52-88.
- Seidenberg, M.S., Waters, G.S., Sanders, M., & Langer, P. (1984) Pre- and post lexical loci of contextual effects on word recognition. *Memory and Cognition*, 12, 315-328.
- Shattuck, S.R. (1975) Speech errors and sentence production. Unpublished doctoral dissertation: Massachusetts Institute of Technology.
- Shattuck-Hufnagel, S. (1979). Speech errors as evidence for a serial order mechanism in sentence production. In W.E. Cooper and E.C.T. Walker (Eds.) *Sentence Processing: Psycholinguistic studies presented to Merrill Garrett*. Hillsdale, NJ: Lawrence Erlbaum.
- Shattuck-Hufnagel, S. (1987). The role of word-onset consonants in speech production planning: New evidence from speech error patterns. In Keller, E. and Gopnik, M. (Eds.) *Motor and Sensory Processes of Language*. Hillsdale, NJ: Lawrence Erlbaum.
- Silvast, M. (1991) Aphasia therapy dialogues. *Aphasiology*, 5, 383-390.
- Sparks, R.W., Helm, N.A., & Albert, M.L. (1974) Aphasia rehabilitation resulting from melodic intonation therapy. *Cortex*, 10, 303-316.
- Spencer, K.A., Doyle, P.J., McNeil, M.R., Wambaugh, J.L., Park, G., & Carroll, B. (2000) Examining the facilitative effects of rhyme in a patient with output lexicon damage. *Aphasiology*, 14, 567-584.
- Springer, L., Glindemann, R., Huber, W., & Willmes, K. (1991) How efficacious is PACE-therapy when language systematic training is incorporated? *Aphasiology*, 5, 391-399.
- Stemberger, J.P. (1983) *Speech Errors and Theoretical Phonology: A review*. Bloomington IN: Indiana University Linguistics Club.
- Stemberger, J.P. (1985). An interactive activation model of language production. In A.W. Ellis (Ed.) *Progress in the psychology of language: Volume 1*. London: Lawrence Erlbaum.
- Swinburn, K., Howard, D. & Porter, G. (manuscript in preparation) CAT Comprehensive Aphasia Test.
- Tabachnik, B.G., & Fidell, L.S. (2001) *Using Multivariate Statistics (Fourth Edition)*. Boston: Allyn and Bacon.
- Tsang, H.L., & Lee, T.M.C. (2003) The effect of ageing on confrontational naming ability. *Archives of Clinical Neuropsychology*, 18, 81-89.

- Ulatowska, H.K., Allard, L., Reyes, B.A., Ford, J., & Chapman, S. (1992) Conversational discourse in aphasia. *Aphasiology*, 6, 325-331.
- Van Gorp, W., Satz, P., Kiersch, M.E., & Henry, R. (1986) Normative data on the Boston Naming Test for a group of normal older adults. *Journal of Clinical and Experimental Neuropsychology*, 8, 702-705.
- Van Turenout, M., Hagoort, P., & Brown, C.M. (1997) Electrophysiological evidence on the time course of semantic and phonological processes in speech production. *Journal of Experimental Psychology: Learning Memory and Cognition*, 23, 787-806.
- Van Turenout, M., Hagoort, P., & Brown, C.M. (1998) Brain activity during speaking: From syntax to phonology in 40 msec. *Science*, 280, 572-574.
- Vigliocco, G., Antonini, T. & Garrett, M.F. (1997). Grammatical gender is on the tip of Italian tongues. *Psychological Science*, 8, 314-317.
- Vigliocco, G., Vinson, D.P., Martin, R.C., & Garrett, M.F. (1999) Is "count" and "mass" information available when the noun is not? An investigation of tip of the tongue states and anomia. *Journal of Memory and Language*, 40, 534-558.
- Warrington, E.K. (1975) The selective impairment of semantic memory. *Quarterly Journal of Experimental Psychology*, 27, 635-657.
- Watson, C.M., Chenery, H.J., & Carter, M.S. (1999) An analysis of trouble and repair in the natural conversations of people with dementia of the Alzheimer's type. *Aphasiology*, 13, 195-218.
- Wheeldon, L., & Monsell, S. (1994) Inhibition of spoken word production by priming a semantic competitor. *Journal of Memory and Language*, 33, 332-336.
- Wheeldon, L.R. & Monsell, S. (1992) The locus of priming effects in spoken word production. *Quarterly Journal of Experimental Psychology*, 44A, 723-761.
- Wheeldon, L.R., & Levelt, W.J.M. (1994) Monitoring the time course of phonological encoding. *Journal of Memory and Language*, 34, 311-334.
- Wilkinson, R. (1995) Aphasia: Conversation analysis of a non-fluent aphasic person. In M. Perkins and S. Howard (Eds.) *Case Studies in Clinical Linguistics*. London: Whurr.
- Wingfield, A. (1967) Perceptual and response hierarchies in object identification. *Acta Psychologica*, 26, 216-226.
- Wingfield, A. (1968) Effects of frequency on identification and naming of objects. *American Journal of Psychology*, 81, 226-234.

APPENDIX ONE

Notations used in conversation analysis

- =** indicates continuous speech between two speakers
- [** indicates overlapping turns
- |**
- .** indicates a micro pause
- (3.5)** indicates a pause and the number of seconds duration

APPENDIX TWO

Conversation Measure: Draft One

Five minute midstream measure of interaction between a person with aphasia and a conversation partner. PA = person with aphasia

A	TOPICS
A1	Total number of topics
A2	Topics initiated by the person with aphasia
A3	Topics initiated by the conversation partner
B	NUMBER OF SPEECH UNITS (=words and other tokens)
B1	Person with aphasia
B2	Conversational partner
C	TURN TAKING
C1	Number of turns: Person with aphasia
C2	Number of turns: Conversational partner
C3	Number of substantive turns: Person with aphasia
C4	Number of substantive turns: Conversational partner
C5	Number of minimal turns: Person with aphasia
C6	Number of minimal turns: Conversational partner
D	WORD RETRIEVAL AND SPEECH ERRORS (PA only)
D1	Total number of content words (excluding paraphasias)
D2	Total number of nouns (subset of content words)
D3	Number of semantic paraphasias/circumlocutions
D4	Number of phonological errors (words/non words: target apparent)
D5	Number of neologisms (target not apparent)
D6	Overuse of pronouns/proforms
D7	Comments (e.g. what's it called?)
D8	Repetition of immediately preceding utterance
D10	Number of pauses greater than 2 seconds within PA's turn
D11	Number of filled pauses within PA's turn (e.g. err, uhm)
E	REPAIR
E1	Number of instances of collaborative repair
E2	Number of turns spent on repair
E3	Number of successful repairs
E4	Number of unsuccessful repairs
	<i>Types of successful repair</i>
E5	Person with aphasia asked for help with word-finding
E6	Conversation partner asked for clarification
E7	Conversation partner corrected person with aphasia
E8	Other
	<i>Types of unsuccessful repair</i>
E9	Conversation partner's unsuccessful guess
E10	Conversation partner fails to get clarification
E11	Other

APPENDIX THREE

Data used in the analysis of inter-rater and intra-rater reliability. Five conversation dyads, conversation at assessment one, rated by two independent raters.

	HM		BB		DA		PH		SC	
	R1	R2	R1	R2	R1	R2	R1	R2	R1	R2
A TOPICS										
A1 Number of topics covered	6	6	11	13	12	6	20	17	1	1
A2 Initiated by person with aphasia	1	1	7	8	3	1	13	9	1	1
A3 Initiated by conversation partner	4	4	4	4	9	5	7	8	0	0
B SPEECH UNITS										
B1 No speech units: person with aphasia	155	153	374	384	288	302	566	547	391	401
B2 No speech units: conversation partner	271	274	369	369	538	541	302	321	328	323
C TURN TAKING										
C1 Turns: person with aphasia	58	58	48	49	102	103	72	69	67	66
C2 Turns: conversation partner	57	57	42	44	87	86	69	68	65	66
C3 Substantive turns: person with aphasia	19	19	36	34	27	31	49	44	45	49
C4 Substantive turns: conversation partner	49	55	37	35	71	72	46	44	50	48
C5 Minimal turns: person with aphasia	8	9	5	7	29	33	15	9	5	5
C6 Minimal turns: conversation partner	3	3	4	5	3	5	20	20	12	15
D LEXICAL RETRIEVAL										
D1 Content words	32	25	130	139	29	31	165	137	95	101
D2 Nouns	18	15	45	47	18	14	30	26	29	33
D3 Semantic paraphasias/circumlocutions	3	0	0	2	0	0	6	6	13	16
D4 Phonological paraphasias	4	4	16	7	23	31	9	9	2	4
D5 Neologisms	2	1	8	6	10	14	4	6	14	9
D6 Overuse of pronouns/proforms	0	0	2	0	1	2	9	7	4	10
D7 Comments	2	1	0	0	0	0	5	5	8	9
D8 Repetition	2	2	17	23	11	10	21	13	4	3
D10 Pause greater than two seconds	19	19	8	7	7	7	3	3	2	2
D11 Filled pause	17	17	33	33	6	3	7	5	5	6
E REPAIR										
E1 Number of repairs	10	12	7	9	15	12	7	8	1	1
E2 Number of turns spent on repair	67	75	28	24	83	80	21	23	121	114
E3 Number of successful repairs	9	10	7	6	14	12	7	8	1	1
E4 No unsuccessful repairs	1	2	0	3	0	0	1	0	0	0
<i>Types of successful repair</i>										
E5 Person with aphasia asked for help	2	3	0	0	0	0	0	3	0	0
E6 Partner asked for clarification	6	5	5	3	0	0	3	0	1	1
E7 Partner corrected person with aphasia	0	0	0	1	0	2	1	1	0	0
E8 Other	1	2	2	2	14	10	3	4	0	0

Types of unsuccessful repair

E9 Partner's unsuccessful guess	1	1	0	0	0	0	0	0	0	0
E10 Partner didn't get clarification	0	0	0	1	0	0	0	0	0	0
E11 Other	0	1	0	2	0	0	0	0	0	0

APPENDIX FOUR

Conversation Measure: Draft Two

Five minute midstream measure of interaction between a person with aphasia and a conversation partner.

B	NUMBER OF SPEECH UNITS (=words and other tokens)
B1	Person with aphasia
B2	Conversational partner
C	TURN TAKING
C1	Number of turns: Person with aphasia
C2	Number of turns: Conversational partner
C3	Number of substantive turns: Person with aphasia
C4	Number of substantive turns: Conversational partner
C5	Number of minimal turns: Person with aphasia
C6	Number of minimal turns: Conversational partner
D	WORD RETRIEVAL AND SPEECH ERRORS (PA only)
D1	Total number of content words (excluding paraphasias)
D2	Total number of nouns (subset of content words)
D4	Number of phonological errors (words/non words: target apparent)
D5	Number of neologisms (target not apparent)
D10	Number of pauses greater than 2 seconds within PA's turn
D11	Number of filled pauses within PA's turn (e.g. err, uhm)
E	REPAIR
E1	Number of instances of collaborative repair
E2	Number of turns spent on repair
E3	Number of successful repairs
E4	Number of unsuccessful repairs

APPENDIX FIVE

DATA USED IN THE ANALYSIS OF TEST-RETEST STABILITY (N= 10 PEOPLE WITH APHASIA)

	HM1	HM2	BB1	BB2	DA1	DA2	PH1	PH2	SC1	SC2	NK1	NK2	OL1	OL2	IK1	IK2	KR1	KR2	QP1	QP2
B SPEECH UNITS																				
B1 Number of speech units: person with aphasia	155	172	384	343	288	123	566	544	401	478	432	469	305	176	125	90	246	299	310	218
B2 Number of speech units: conversation partner	271	303	369	269	538	496	302	284	323	201	118	232	334	542	229	246	340	350	367	374
C TURNS																				
C1 Total number of turns: person with aphasia	58	50	49	35	102	51	72	71	66	51	24	38	36	38	43	42	69	54	53	47
C2 Total number of turns conversation partner	57	47	44	31	87	47	69	66	66	46	23	38	37	37	42	42	68	55	52	49
C3 Substantive turns: person with aphasia	19	21	34	18	27	13	49	42	49	32	24	30	28	15	20	17	46	32	45	29
C4 Substantive turns: conversation partner	49	42	35	27	71	52	46	39	48	23	18	23	30	30	37	32	46	42	45	41
C5 Minimal turns: person with aphasia	8	13	7	12	29	26	15	13	5	12	0	4	3	19	8	8	13	20	5	12
C6 Minimal turns: conversation partner	3	2	5	3	3	0	20	21	15	19	5	10	6	0	2	4	22	11	6	7
D LEXICAL RETRIEVAL																				
D1 Content words	32	32	139	91	29	11	165	141	101	125	130	145	95	41	21	12	76	98	106	62
D2 Nouns	18	20	47	49	18	9	30	50	33	46	35	27	19	11	6	6	46	32	29	16
D4 Phonological paraphasias	4	6	7	1	23	2	9	7	4	7	8	3	5	1	6	8	18	2	23	22
D5 Neologisms	2	7	6	1	10	1	4	9	9	9	0	0	3	3	1	8	4	8	9	7
D10 Within turn pause	12	12	4	1	3	15	3	1	1	6	17	2	3	2	5	4	3	2	1	1
D11 Filled pause	17	9	33	32	6	13	7	15	6	3	15	23	9	2	20	12	37	37	16	9
E REPAIR																				
E1 Number of instances of repair	10	6	9	5	15	11	7	7	1	4	5	7	4	1	7	9	8	4	7	9
E2 Number of turns spent on repair	67	66	24	17	83	46	21	38	114	36	30	32	12	3	62	53	49	21	40	48
E3 Successful repairs	9	3	6	5	14	11	7	5	1	3	5	7	4	1	5	9	8	4	7	9
E4 Unsuccessful repairs	1	3	3	0	0	0	1	2	0	1	0	0	0	0	2	0	0	0	0	0
F PROPORTIONAL DATA																				
F1 Substantive turns/turns	0.33	0.42	0.69	0.51	0.26	0.25	0.68	0.59	0.74	0.63	1	0.79	0.78	0.39	0.47	0.4	0.67	0.59	0.85	0.62
F2 Minimal turns / turns	0.14	0.26	0.14	0.34	0.28	0.51	0.21	0.18	0.08	0.24	0	0.11	0.08	0.5	0.19	0.19	0.19	0.37	0.09	0.26
F3 Content words /speech units	0.21	0.19	0.36	0.27	0.1	0.09	0.29	0.26	0.25	0.26	0.3	0.31	0.31	0.23	0.17	0.13	0.31	0.33	0.34	0.28
F4 Nouns / speech units	0.12	0.12	0.12	0.14	0.06	0.07	0.05	0.09	0.08	0.1	0.08	0.06	0.06	0.06	0.05	0.07	0.19	0.11	0.09	0.07
F5 Filled pauses /speech units	0.11	0.05	0.09	0.09	0.02	0.11	0.01	0.03	0.01	0.01	0.03	0.05	0.03	0.01	0.16	0.13	0.15	0.12	0.05	0.04
F6 Instances of repair/ turns	0.09	0.06	0.10	0.08	0.08	0.11	0.05	0.05	0.01	0.04	0.11	0.09	0.05	0.01	0.08	0.11	0.06	0.04	0.07	0.09

APPENDIX SIX

Conversation Measure: Draft Three

Five minute midstream measure of interaction between a person with aphasia and a conversation partner.

B	NUMBER OF SPEECH UNITS (=words and other tokens)
B1	Person with aphasia
B2	Conversational partner
C	TURN TAKING
C1	Number of turns: Person with aphasia
C2	Number of turns: Conversational partner
C3	Number of substantive turns: Person with aphasia
C4	Number of substantive turns: Conversational partner
C5	Number of minimal turns: Person with aphasia
C6	Number of minimal turns: Conversational partner
D	WORD RETRIEVAL AND SPEECH ERRORS (PA only)
D1	Total number of content words (excluding paraphasias)
D2	Total number of nouns (subset of content words)
D11	Number of filled pauses within PA's turn (e.g. err, uhm)
E	REPAIR
E1	Total instances of repair
F	PROPORTIONAL DATA
F1	Substantive turns / turns (PA)
F2	Minimal turns / turns (PA)
F3	Content words / speech units
F4	Nouns / speech units
F5	Filled pauses / speech units
F6	Instances of repair / total turns